

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年5月22日(22.05.2020)



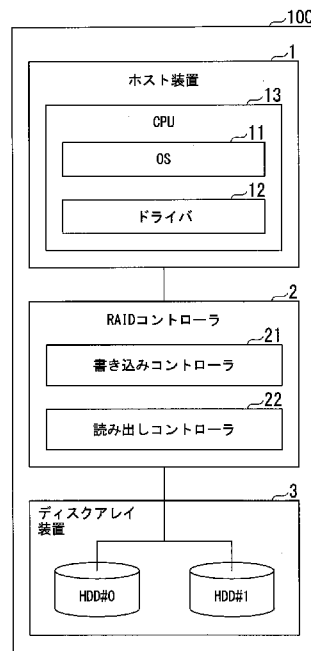
(10) 国際公開番号

WO 2020/100883 A1

- (51) 国際特許分類:
G06F 11/20 (2006.01) *G06F 3/06* (2006.01)
- (21) 国際出願番号: PCT/JP2019/044304
- (22) 国際出願日: 2019年11月12日(12.11.2019)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2018-215857 2018年11月16日(16.11.2018) JP
- (71) 出願人: NECプラットフォームズ株式会社
(**NEC PLATFORMS, LTD.**) [JP/JP]; 〒2138511
神奈川県川崎市高津区北見方二丁目
6番1号 Kanagawa (JP).
- (72) 発明者: 高沢 和義 (**TAKAZAWA Kazuyoshi**);
〒2138511 神奈川県川崎市高津区北見方二
丁目6番1号 NECプラットフォーム
ズ株式会社内 Kanagawa (JP).
- (74) 代理人: 棚井 澄雄, 外 (**TANAI Sumio et al.**);
〒1006620 東京都千代田区丸の内一丁
目9番2号 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保
護が可能): AE, AG, AL, AM, AO, AT, AU, AZ,
BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH,
CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,

(54) **Title:** DISK ARRAY SYSTEM, CONTROL DEVICE, AND DATA INPUT/OUTPUT METHOD

(54) 発明の名称: ディスクアレイシステム、制御装置、およびデータ入出力方法



- 1 Host device
2 RAID controller
3 Disk array device
12 Driver
21 Writing controller
22 Reading controller

(57) **Abstract:** A control device employed in a disk array system comprising a plurality of hard disks, said control device being provided with a writing unit that writes data to be written comprising a plurality of unit data to a plurality of hard disks in a distributed manner, and a reading unit that simultaneously reads a plurality of unit data in parallel from the plurality of hard disks. The reading unit simultaneously reads, from a prescribed sector in the plurality of hard disks, a contiguous unit data group in an array of a plurality of unit data, and the next unit data group is read simultaneously from a sector different from the prescribed sector. When an abnormality is detected in unit data stored in one sector of one hard

MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類：

- 一 国際調査報告（条約第21条(3)）

disk among the plurality of hard disks, the read sector of the unit data is set to the one sector of another hard disk, and the read sector for other unit data is set to another sector of the one hard disk.

(57) 要約：複数のハードディスクを備えたディスクアレイシステムに適用される制御装置であって、複数の単位データからなる書込対象データを複数のハードディスクに分散して書き込む書込部と、複数のハードディスクから並行して複数の単位データを同時に読み出す読出部と、を備える。読出部は、複数の単位データの配列において連続する単位データ群を複数のハードディスクにおける所定のセクタから同時に読み出すとともに、次の単位データ群は前記所定のセクタと異なるセクタから同時に読み出す。複数のハードディスクのうちのハードディスクの一のセクタに記憶された単位データに異常を検出すると、単位データの読み出しセクタを他のハードディスクの一のセクタに設定し、他の単位データを読み出すセクタを一のハードディスクの他のセクタに設定する。

明 細 書

発明の名称：

ディスクアレイシステム、制御装置、およびデータ入出力方法

技術分野

[0001] 本発明は、ディスクアレイシステム、制御装置、およびデータ入出力方法に関する。

背景技術

[0002] ハードディスクドライブ（HDD）の可用性や入出力性能を向上するための技術として、RAID（Redundant Array of Inexpensive Disks）が開発されている。例えば、ストライピング方式のRAID0では、複数のディスクにデータを分散して書き込むことにより、書き込み速度を向上することができる。また、ミラー方式のRAID1では、複数のディスクに同じデータを書き込むことにより可用性を向上させている。

[0003] 特許文献1は、複数のディスクでミラーリング構成したRAID1において、読み込むデータのサイズによらずに、ディスクからの読み取り速度を高速化する情報処理装置を開示している。つまり、2つのハードディスクドライブ（HDD）でRAID1を構成した場合、所定サイズより大きなデータを読み出すときに、前半のデータを一方のHDDから読み出すとともに、後半のデータを他方のHDDから同時に読み出すことにより、読み出し速度を向上した技術を開示している。特許文献2は、複数のハードディスクドライブ（HDD）でミラーリング構成したRAID1の記憶媒体において、データの読み出し速度を高速化した技術を開示している。記憶装置のRAIDコントローラには論理アドレスと2台のハードディスクドライブ（HDD）のセクタアドレスとの対応付けが格納されており、複数のセクタにより構成されるブロック単位に所定数のブロックを読み出し単位としてデータを並列に読み出すことで、読み出し処理の高速化を実現している。

先行技術文献

特許文献

[0004] 特許文献1：特開2017-16579号公報

特許文献2：特開2012-103856号公報

発明の概要

発明が解決しようとする課題

[0005] 特許文献1に係るミラーリング方式のRAID1では、所定サイズのデータを読み出す場合には、2台のHDDのうち一方のHDDから読み出すため、可用性を向上するものの、記憶装置全体としての読み出し性能は向上しない。また、RAID0構成では、複数のHDDに別々のデータを書き込むため、読み出し速度は向上するが可用性は向上しない。特許文献2では、RAIDコントローラが2台のHDDの連続ブロックに対して、ブロック毎（2セクタ毎）に並列でアクセスしているため、読み出し速度は向上するが、可用性は向上しない。

課題を解決するための手段

[0006] 本発明の第一態様によれば、複数のハードディスクを備えたディスクアレイシステムに適用される制御装置は、複数の単位データからなる書込対象データを、複数のハードディスクに分散して書き込む書込部と、複数のハードディスクから並行して複数の単位データを同時に読み出す読出部と、を備える。書込部は、複数の単位データの配列を変更せずに複数のハードディスク内の複数のセクタに分散して書き込み、読出部は、複数の単位データの配列において連続する単位データ群を複数のハードディスクにおける所定のセクタから同時に読み出すとともに、次の単位データ群は所定のセクタと異なるセクタから同時に読み出す。複数のハードディスクのうちのハードディスクの一のセクタに記憶された単位データに異常を検出すると、単位データの読み出しセクタを他のハードディスクの一のセクタに設定し、他の単位データを読み出すセクタを一のハードディスクの他のセクタに設定する。或いは

、書込部は、読出部が複数のハードディスクの複数のセクタにおいて所定セクタ順序で複数の単位データの配列において連続する単位データ群を同時に読み出すことができるように、複数の単位データの配列を変更して、複数のハードディスクの前記複数のセクタに書き込んでもよい。この場合、読出部は、複数のハードディスクから所定のセクタ順序で複数の単位データの配列において連続する単位データ群を同時に読み出す。

[0007] 本発明の第二態様によれば、複数のハードディスクを備えたディスクアレイシステムに適用されるデータ入出力方法において、複数の単位データからなる書込対象データにおける複数の単位データの配列を変更せずに複数のハードディスク内の複数のセクタに分散して書き込み、複数の単位データの配列において連続する単位データ群を複数のハードディスクにおける所定のセクタから同時に読み出すとともに、次の単位データ群は所定のセクタと異なるセクタから同時に読み出す。複数のハードディスクのうちのハードディスクの一のセクタに記憶された単位データに異常を検出すると、単位データの読み出しセクタを他のハードディスクの一のセクタに設定し、他の単位データを読み出すセクタを一のハードディスクの他のセクタに設定する。或いは、複数の単位データからなる書込対象データにおける複数のハードディスクの複数のセクタにおいて所定セクタ順序で複数の単位データの配列において連続する単位データ群を同時に読み出すことができるように、複数の単位データの配列を変更して、複数のハードディスクの複数のセクタに書き込み、複数のハードディスクから所定のセクタ順序で前記複数の単位データの配列において連続する単位データ群を同時に読み出す。

[0008] 本発明の第三態様によれば、ディスクアレイシステムは、複数のハードディスクと、複数のハードディスクに対して複数の単位データの読み出しと書き込みを制御する制御装置と、を備える。制御装置は、複数の単位データの配列を変更せずに前記複数のハードディスク内の複数のセクタに分散して書き込む書込部と、複数の単位データの配列において連続する単位データ群を複数のハードディスクにおける所定のセクタから同時に読み出すとともに、

次の単位データ群は所定のセクタと異なるセクタから同時に読み出す読出部と、を備える。読出部は、複数のハードディスクのうち一のハードディスクの一のセクタに記憶された単位データに異常を検出すると、単位データの読み出しセクタを他のハードディスクの一のセクタに設定し、他の単位データを読み出すセクタを前記一のハードディスクの他のセクタに設定する。或いは、制御装置は、複数のハードディスクの複数のセクタにおいて所定セクタ順序で複数の単位データの配列において連続する単位データ群を同時に読み出すことができるように、複数の単位データの配列を変更して、複数のハードディスクの複数のセクタに書き込む書込部と、複数のハードディスクから所定のセクタ順序で複数の単位データの配列において連続する単位データ群を同時に読み出す読出部と、を備える。

発明の効果

[0009] 本発明によれば、ディスクアレイシステムの可用性や読み出し性能を向上するという効果を奏する。

図面の簡単な説明

[0010] [図1]本発明の実施形態に係る情報処理システムの構成を示すブロック図である。

[図2]本発明の第一実施形態に係るデータ入出力方法を説明する説明図である。

[図3]本発明の第一実施形態に係るデータ入出力方法を説明する説明図である。

[図4]本発明の第一実施形態に係るデータ入出力方法を説明する説明図である。

[図5]本発明の第一実施形態によるハードディスクのリビルド処理を説明する説明図である。

[図6]本発明の第二実施形態に係るデータ入出力方法を説明する説明図である。

[図7]本発明の第二実施形態による複数のHDDのセクタ毎のデータ格納順序

を現す２種類の表を示す。

[図8]本発明の第二実施形態に係るデータ入出力方法を説明する説明図である。

[図9]本発明の第二実施形態に係るデータ入出力方法を説明する説明図である。

[図10]本発明の第二実施形態によるハードディスクのリビルド処理を説明する説明図である。

[図11]本発明の第二実施形態において３台のHDDによるデータの格納順を示す説明図である。

[図12]本発明の実施形態における制御装置の最小構成を示すブロック図である。

発明を実施するための形態

[0011] 本発明に係るディスクアレイシステムの制御装置およびデータ入出力方法について、添付図面を参照して実施例とともに詳細に説明する。複数の図面において、同一の構成部分については同一の符号を付して、その説明を適宜割愛する。

[0012] 図１は、本発明の実施形態に係る情報処理システム１００の構成を示すブロック図である。情報処理システム１００は、ホスト装置１と、RAIDコントローラ２と、ディスクアレイ装置３と、を備える。ホスト装置１は、オペレーティングシステム（OS）１１とドライバ１２を具備するCPU（Central Processing Unit）１３を備えている。OS１１は、ホスト装置１で実行されるアプリケーションなどを発行するIO（Input／Output）命令を実行し、ディスクアレイ装置３への書き込みや読み出しを指示する。ドライバ１２は、RAIDコントローラ２のデバイスドライバである。なお、OS１１とドライバ１２は、ソフトウェアにより実現される。例えば、CPU１３がディスクアレイ装置１３に格納されているプログラムを読み出して実行することにより、OS１１とドライバ１２の機能を実現する。

[0013] R A I Dコントローラ2は、書き込みコントローラ21と、読み出しコントローラ22と、を備える。書き込みコントローラ21と読み出しコントローラ22は、ミラーリング方式のR A I D1に必要な機能を有するソフトウェア（ファームウェア）により実現される。後述するように、書き込みコントローラ21と読み出しコントローラ22は、R A I D1によるデータの冗長化機能に加えて、読出処理を高速化する機能を備えている。R A I Dコントローラ2は、ディスクアレイ装置3に接続されている。ディスクアレイ装置3は、2台のハードディスクドライブHDD#0とHDD#1を備えており、HDD#0とHDD#1はデータの冗長性を保証する1つの論理ディスクを構成している。書き込みコントローラ21と読み出しコントローラ22は、HDD#01とHDD#02への入出力（I O）処理を実行する。

[0014] 次に、情報処理システム100における入出力処理について説明する。ホスト装置1において、OS11がI O命令を発行すると、ドライバ12がそのI O命令をR A I Dコントローラ2に出力して、入出力処理の実行を要求する。I O命令が書き込み処理（W R I T E）の場合、書き込みコントローラ21は、HDD#0とHDD#1に同じデータを書き込む。これにより、R A I D1構成と同様のデータの冗長性が確保される。I O命令が読み出し処理（R E A D）の場合、読み出しコントローラ22は、HDD#0とHDD#1の各々から読み出し対象データの一部ずつを同時に並行して読み出す。2台のHDDのうち一方のHDDから読み出し対象データの全てを読み出すR A I D1構成に比べて、読み出しコントローラ22は、2倍の速度でデータを読み出すことができる。

[0015] [第一実施形態]

次に、本発明の第一実施形態に係る制御装置として、R A I Dコントローラ2の処理について詳細に説明する。図1は、本発明の第一実施形態に係るデータ入出力方法を説明する説明図である。図2において、矢印上の数字列やHDD#0とHDD#1内の数字は、書き込みデータの構成や読み出しデータの構成を示している。例えば、書き込みコントローラ21への入力デー

タ「0 1 2 3 4 5 6 7」は、書き込み対象データが、「データ0」、「データ1」、「データ2」、「データ3」、「データ4」、「データ5」、「データ6」、「データ7」により構成されていることを示す。この「データ0」乃至「データ7」の各々を便宜的に単位データと呼ぶ。図2に例示する書き込み対象データでは、8個の単位データが「データ0」乃至「データ7」の順に配列されている。

[0016] (データ書き込み処理)

ディスクアレイ装置3への8セクタ長のデータ(データ0～データ7)の書き込み要求があると、書き込みコントローラ21は、HDD#0とHDD#1の両方に、8セクタ分のデータ書き込み処理を実行する。書き込みコントローラ21は、書き込み対象データ(データ0～データ7)を取得すると、この配列の状態で書き込み対象データをHDD#0とHDD#1のそれぞれに書き込む。このとき、書き込みコントローラ21は、HDD#0とHDD#1の同じセクタ位置にそれぞれの単位データを順に書き込む。図2の表は、HDD#0とHDD#1におけるセクタ位置と、書き込みコントローラ21により書き込まれる単位データとの関係を示す。図2の表に示すように、HDD#0とHDD#1の両方において、セクタ0にはデータ0が格納され、セクタ1にはデータ1が格納される。つまり、セクタNにはデータN(N=0～7)が格納される。第一実施形態では、書き込みコントローラ21の動作は、RAID1構成に対する書き込み処理と同一である。

[0017] (通常時のデータ読み出し処理)

次に、通常時(非障害時)のデータ読み出し処理について説明する。RAIDコントローラ2がホスト装置1から8セクタ長のデータ読み出し指示を受けると、読み出しコントローラ22は、HDD#0とHDD#1の異なるセクタから並行して単位データの読み出しを行う。例えば、読み出しコントローラ22は、HDD#0の偶数セクタ(セクタ0、2、4、6)から偶数の符号が付された単位データ(データ0、2、4、6)を読み出すとともに、HDD#1の奇数セクタ(セクタ1、3、5、7)から奇数の符号が付さ

れた単位データ（データ 1、3、5、7）を同時に読み出す。具体的には、読み出しコントローラ 22 は、HDD # 0 からセクタ 0、セクタ 2、セクタ 4、セクタ 6 のデータを順に読み出し、HDD # 1 からセクタ 1、セクタ 3、セクタ 5、セクタ 7 のデータを順に読み出す。このとき、読み出しコントローラ 22 は、書き込み対象データにおける単位データの配列において隣り合う（又は、連続する）単位データを同時に読み出す。具体的には、読み出しコントローラ 22 は、HDD # 0 のセクタ 0 と HDD # 1 のセクタ 1 から 2 個の単位データ（データ 0 とデータ 1）を同時に読み出す。そして、読み出しコントローラ 22 は、2 個の読み出した単位データ（データ 0 とデータ 1）を同時に読み出して読み出し要求元であるホスト装置 1 へ送信する。

[0018] 次に、読み出しコントローラ 22 は、HDD # 0 のセクタ 2 と HDD # 1 のセクタ 3 からデータ 2 とデータ 3 を同時に読み出して、ホスト装置 1 へ同時に送信する。同様に、読み出しコントローラ 22 は、HDD # 0 のセクタ 4 と HDD # 1 のセクタ 5 からデータ 4 とデータ 5 を同時に読み出してホスト装置 1 へ同時に送信し、最後に、HDD # 0 のセクタ 6 と HDD # 1 のセクタ 7 からデータ 6 とデータ 7 を同時に読み出してホスト装置 1 へ同時に送信する。このようにして、HDD # 0 と HDD # 1 から 8 セクタ分の単位データを読み出す。このような処理により、複数のデータの冗長性を確保しつつ、2 台の HDD のうち一方の HDD からのみデータを読み出す場合に比べて、2 倍の速度でデータを読み出すことができる。

[0019] 情報処理システムでは、記憶装置に格納されるデータには空間的局所性（情報処理に必要なデータはアドレスが近接した記憶位置に纏まって記憶される性質）があるため、意味のある順番で単位データが配列されて、書き込みや読み出しが行われる。上記の読み出しコントローラ 22 による読み出し処理では、隣接する単位データ（例えば、データ 0 とデータ 1、データ 2 とデータ 3、など）を同時に取得することができるので、データ読み出し後の OS 11 の処理においても高速化が期待できる。例えば、8 セクタ長のデータを前半データ（データ 0 乃至データ 3）と後半データ（データ 4 乃至データ

7) に分割し、前半データと後半データを1セクタずつ同時に並行して読み出す方法が提案されている（例えば、特許文献1）。この読み出し方法では、前半データの先頭のデータ0と後半データの先頭のデータ4を同時に読み出して読み出し要求元へ送信し、次に、データ1とデータ5を読み出して要求元へ送信するような手順を繰り返す。この読み出し方法では、読み出し要求元側の処理においてデータ0とデータ1が必要な場合、セクタ1の後のセクタ1からデータ1を読み出すまでの待ち時間が発生する。これに対して、第一実施形態では、連続する単位データを高速に読み出すことができるため、連続する単位データを読み出すときに発生する待ち時間を低減し、読み出し処理を高速化することができる。

[0020] （データ異常検出時の読み出し処理）

次に、ある単位データ（以下の例では、HDD#0のセクタ2に格納されているデータ2）が壊れており、正常にデータを読み出すことができない場合（データ異常検出時）の動作について説明する。図3は、本発明の第一実施形態に係るデータ入出力方法を説明する説明図である。例えば、8セクタ分のデータの読み出し要求があると、通常時の読み出し処理では、読み出しコントローラ22は、HDD#0のセクタ0とHDD#1のセクタ1からデータ0とデータ1を同時に読み出して、読み出し要求元へ同時に送信する。次に、読み出しコントローラ22は、HDD#0のセクタ2とHDD#1のセクタ3からデータ2とデータ3を同時に読み出すことになるが、データ2が壊れていることを検出すると、読み出しコントローラ22は、データ2とデータ3を一旦破棄する。その後、読み出しコントローラ22は、異常が検出されたデータ2の読み出し先をHDD#0のセクタ2からHDD#1のセクタ2に切り替えるとともに、HDD#0の読み出し対象セクタをセクタ2から1つ先にずらしてセクタ3に設定する。つまり、読み出しコントローラ22は、HDD#0のセクタ2に格納されたデータ2の異常を検出すると、それ以降の単位データの読み出しにおいて、HDD#0の読み出し対象セクタを奇数セクタに変更し、HDD#1の読み出し対象セクタを偶数セクタに

変更する。これにより、読み出しコントローラ 22 は、HDD # 1 のセクタ 1 からデータ 1 を読み出した後、HDD # 1 のセクタ 2 と HDD # 0 のセクタ 3 からデータ 2 とデータ 3 を同時に読み出して、読み出し要求元へ同時に送信する。次に、読み出しコントローラ 22 は、HDD # 1 のセクタ 4 と HDD # 0 のセクタ 5 からデータ 4 とデータ 5 を同時に読み出して、読み出し要求元へ同時に送信する。このようにして、読み出しコントローラ 22 は、HDD # 0 からデータ 0、データ 3、データ 5、データ 7 を読み出し、HDD # 1 からデータ 1、データ 2、データ 4、データ 6 を読み出す。

[0021] 上記の読み出し制御により、読み出しコントローラ 22 は、HDD # 0 と HDD # 1 から読み出したデータ 2 とデータ 3 を一度破棄して、HDD # 0 と HDD # 1 の読み出しセクタを変更して、データ 2 とデータ 3 を読み出しなおすことにより、1 セクタの単位データ読み出し分の遅れは生じるものの、HDD # 0 と HDD # 1 から同時に連続する単位データを順次読み出すことができるため、データ異常検出時においても通常時と同等な高速のデータ読み出し処理を実現することができる。

[0022] (HDD 障害時)

次に、HDD の一部のセクタの異常ではなく、HDD 全体（以下の例では、HDD # 0）が固定的に故障した場合の動作について説明する。図 4 は、本発明の第一実施形態に係るデータ入出力方法を説明する説明図である。HDD # 0 が故障している場合、RAID コントローラ 2 は、通常の RAID 1 と同様に、HDD # 0 を閉塞する。そして、8 セクタ分のデータの読み出し要求があると、読み出しコントローラ 22 は、HDD # 1 からのみ全ての単位データの読み出しを行う。つまり、読み出しコントローラ 22 は、HDD # 1 のセクタ 0 乃至セクタ 7 から順に単位データを読み出して、1 セクタの単位データを読み出す度に、読み出し要求元（ホスト装置 1）へデータ 0 乃至データ 7 を送信する。なお、単位データの書き込みについては、書き込みコントローラ 21 が、閉塞されていない HDD # 1 に対して図 2 の表に示したようにデータ 0 乃至データ 7 を書き込む。また、HDD 障害時の読み出

し処理と書き込み処理はRAID1と同様である。

[0023] (リビルド処理)

図5は、本発明の第一実施形態によるハードディスクのリビルド処理を説明する説明図である。HDD#0が故障した場合、HDD#0を交換して、新たなHDD#0にHDD#1に書き込まれたデータと同じデータを書き込む。なお、新たなHDD#0にデータを書き込んで新たなHDD#0とHDD#1によるRAID1の状態に再構成する処理をリビルド処理という。具体的には、読み出しコントローラ22は、HDD#1からセクタ順に単位データを読み出して、書き込みコントローラ21へ送る。書き込みコントローラ21は、データが書き込まれていないHDD#0の各セクタへHDD#1の同じセクタに書き込まれた単位データを書き込む。ハードディスクのリビルド処理が終了すると、HDD#0とHDD#1の両者に同一の単位データが書き込まれたことになる。リビルド処理後は、図2を用いて説明したように、書き込み時には、HDD#0とHDD#1の同じセクタに同じ単位データを書き込み、読み出し時には、HDD#0の偶数セクタとHDD#1の奇数セクタから同時に単位データを読み出すような制御を行う。これにより、第一実施形態によるディスクアレイ装置3の冗長性の確保と、高速読み出し処理を実現することができる。

[0024] 上述のように、第一実施形態によれば、ディスクアレイ装置に備わる複数のHDDに同一のデータを書き込むことで冗長性を保ちつつ、複数のHDDから同時に別の単位データを読み出すことで、読み出し処理の高速化を図ることができる。特に、第一実施形態では、複数のHDDから連続する単位データを同時に読み出すことができるので、データ読み出し後の処理の高速化を図ることができる。なお、第一実施形態では、ディスクアレイ装置を2台のHDDで構成する例を示したが、HDDの数を3台以上としてもよい。その場合、読み出しコントローラ22は、複数のHDDのセクタについて1ずつずらしたセクタから同時に単位データを読み出す制御を行う。

[0025] [第二実施形態]

第一実施形態では、読み出しコントローラ 22 は、HDD # 0 の偶数セクタと HDD # 1 の奇数セクタから同時に単位データを読み出す制御について説明した。第二実施形態では、HDD # 0 と HDD # 1 の同一セクタから同時に単位データを読み出す制御について説明する。図 6 は、本発明の第二実施形態に係るデータ入出力方法を説明する説明図である。図 7 は、本発明の第二実施形態による 2 台の HDD のセクタ毎のデータ格納順序を現す 2 種類の表を示す。第二実施形態においても、図 1 の情報処理システム 100 の構成に基づいて説明を行う。但し、第二実施形態では、第一実施例の書き込みコントローラ 21 と読み出しコントローラ 22 と異なる機能を発揮しているので、図 8 乃至図 10 に示すように、書き込みコントローラ 21 a と読み出しコントローラ 22 a と表記する。

[0026] (データの書き込み処理)

ディスクアレイ装置 3 への 8 セクタ長（データ 0 乃至データ 7）の書き込み要求があると、書き込みコントローラ 21 a は、HDD # 0 と HDD # 1 の両方に 8 セクタ分の単位データを書き込む。第二実施形態では、書き込みコントローラ 21 a は、HDD # 0 と HDD # 1 に書き込む単位データの順序を予め入れ替える。つまり、読み出しコントローラ 22 a が HDD # 0 と HDD # 1 における同一のセクタを始点として単位データをセクタ順に読み出した場合、書き込み前の元の単位データの配列（データ 0 乃至データ 7）において隣接する単位データを同時に読み出すことができるように、HDD # 0 と HDD # 1 に書き込む単位データの順序を予め入れ替える。例えば、書き込みコントローラ 21 a が複数の単位データを HDD # 0 に格納する場合、偶数の符号が付された単位データの並び順を維持したまま前半データとして配列し、奇数の符号が付された単位データの並び順を維持したまま後半データとして配列する。具体的には、元の単位データの配列（データ 0 乃至データ 7）を、前半データ（データ 0、データ 2、データ 4、データ 6）と後半データ（データ 1、データ 3、データ 5、データ 7）に並び替えて、前半データと後半データをセクタ順に HDD # 0 に格納する（図 7（a）の表の

上段参照)。

[0027] 一方、書き込みコントローラ 21a が複数の単位データを HDD # 1 に格納する場合、奇数の符号が付された単位データの並び順を維持したまま前半データとして配列し、偶数の符号が付された単位データの並び順を維持したまま後半データとして配列する。具体的には、元の単位データの配列(データ 0 乃至データ 7)を、前半データ(データ 1、データ 3、データ 5、データ 7)と後半データ(データ 0、データ 2、データ 4、データ 6)に並び替えて、前半データと後半データをセクタ順に HDD # 1 に格納する(図 7 (a)の表の下段参照)。

[0028] 図 6 は、上記の書き込み処理を示しており、図 7 (a)の表は、8セクタ長のデータを HDD # 0 と HDD # 1 に書き込んだ状態を示している。すなわち、HDD # 0 のセクタ 0 にはデータ 0 が格納され、HDD # 1 のセクタ 0 にはデータ 1 が格納される。また、HDD # 0 セクタ 1 にはデータ 2 が格納され、HDD # 1 のセクタ 1 にはデータ 3 が格納される。このようにして、HDD # 0 のセクタ 0 乃至セクタ 7 にはデータ 0、データ 2、データ 4、データ 6、データ 1、データ 3、データ 5、データ 7 が格納され、HDD # 1 のセクタ 0 乃至セクタ 7 にはデータ 1、データ 3、データ 5、データ 7、データ 0、データ 2、データ 4、データ 6 が格納される。これにより、HDD # 0 と HDD # 1 の同一セクタから同時に単位データを読み取ると、元のデータ配列(データ 0 乃至データ 7)において連続する単位データを取得することができる。このように、第二実施形態では、読み出しコントローラ 22a が、HDD # 0 と HDD # 1 のセクタ 0 から順に単位データを読み出したときに連続する単位データが同時に得られるように元のデータ配列を並び替えて、複数の単位データを HDD # 0 と HDD # 1 にそれぞれ書き込む。また、HDD # 0 と HDD # 1 に書き込むセクタ長により書き込みセクタは異なるため、読み出しコントローラ 22a は、セクタ長の情報も同時に(書き込み開始セクタと対応付けて) HDD # 0 と HDD # 1 に書き込んで保持する。図 7 (b)の表は、HDD # 0 と HDD # 1 に 4セクタ長の単位データ

書き込む場合の、単位データ書き込みセクタを示す。

[0029] (通常時の読み出し処理)

次に、本発明の第二実施形態に係る読み出し処理について説明する。まず、通常時（非障害時）の動作について説明する。ホスト装置 1 から 8 セクタ長のデータ読み出し指示を受けると、読み出しコントローラ 22 a は、HDD # 0 と HDD # 1 の同一セクタから並行して同時に単位データを読み出す。図 7 (a) の表では、読み出しコントローラ 22 は、まず、HDD # 0 と HDD # 1 のセクタ 0 から同時に単位データを読み出す。読み出しコントローラ 22 a は、HDD # 0 と HDD # 1 のセクタ 0 から読み出した単位データ（データ 0 とデータ 1）を同時に読み出し要求元（ホスト装置 1）へ送信する。

[0030] 次に、読み出しコントローラ 22 a は、HDD # 0 と HDD # 1 のセクタ 1 からデータ 2 とデータ 3 を同時に読み出して、読み出し要求元へ送信する。また、読み出しコントローラ 22 a は、HDD # 0 と HDD # 1 のセクタ 2 からデータ 4 とデータ 5 を同時に読み出して、読み出し要求元へ送信する。最後に、読み出しコントローラ 22 a は、HDD # 0 と HDD # 1 のセクタ 3 からデータ 6 とデータ 7 を同時に読み出して、読み出し要求元へ送信する。このようにして、読み出しコントローラ 22 a は、HDD # 0 と HDD # 1 から 8 セクタ分の単位データ（データ 0 乃至データ 7）を読み出す。上記の読み出し処理により、第二実施形態では、第一実施形態と同様に、HDD に格納されるデータの冗長性を確保しつつ、2 台の HDD のうち一方の HDD からのみデータを読み出す場合に比べて、2 倍の速度でデータを読み出すことができる。

[0031] ここで、第一実施形態と第二実施形態に係る読み出し処理について比較する。第一実施形態では、HDD # 0 の偶数セクタと HDD # 1 の奇数セクタから同時に単位データを読み出している。1 つの HDD に着目すると、セクタを 1 つ飛ばして単為データを読み込むことになる。HDD の性質として、不連続なセクタからデータを読み出す場合には、セクタを再度サーチする処

理が行われる。この観点から、第一実施形態の読み出し処理には、再度サーチする必要性から若干読み出し速度の低下が発生する。一方、第二実施形態では、読み出しコントローラ 22a は、HDD # 0 と HDD # 1 において連続するセクタから順次データを読み出している。そのため、HDD について不連続なセクタを再度サーチする必要がなく、高速な読み出し処理が可能となる。

[0032] (データ異常検出時の読み出し処理)

次に、ある単位データ（以下の例では、HDD # 0 のセクタ 1 に格納されているデータ 2）が壊れており正常に読み出すことができなかった場合の動作について説明する。図 8 は、本発明の第二実施形態に係るデータ入出力方法を説明する説明図である。読み出し要求元から 8 セクタ分のデータの読み出し要求があると、通常の読み出し処理では、読み出しコントローラ 22a は、まず、HDD # 0 のセクタ 0 と HDD # 1 のセクタ 0 からデータ 0 とデータ 1 を同時に読み出して読み出し要求元へ同時に送信する。次に、読み出しコントローラ 22a は、HDD # 0 のセクタ 1 と HDD # 1 のセクタ 1 からデータ 2 とデータ 3 を同時に読み出したときにデータ 2 が壊れていることを検出すると、読み出しコントローラ 22a は、データ 2 とデータ 3 を一旦破棄する。その後、読み出しコントローラ 22a は、異常が検出されたデータ 2 の読み出し先を HDD # 0 から HDD # 1 へ切り替える。具体的には、データ 2 の読み出しセクタを、HDD # 0 のセクタ 1 を HDD # 1 のセクタ 5 へ設定する。また、HDD # 0 の読み出しセクタもセクタ 5 に設定する。そして、読み出しコントローラ 22a は、HDD # 0 のセクタ 5 と HDD # 1 のセクタ 5 からデータ 3 とデータ 2 を同時に読み出して読み出し要求元へ同時に送信する。次に、読み出しコントローラ 22a は、HDD # 0 のセクタ 6 と HDD # 1 のセクタ 6 からデータ 5 とデータ 4 を同時に読み出して読み出し要求元に同時に送信する。最後に、読み出しコントローラ 22a は、HDD # 0 のセクタ 7 と HDD # 1 のセクタ 7 からデータ 7 とデータ 6 を同時に読み出して読み出し要求元へ同時に送信する。

[0033] 上述のように、読み出しコントローラ 22 a は、異常データの検出前は、HDD # 0 から偶数の符号が付された単位データを読み出すとともに、HDD # 1 から奇数の符号が付された単位データを同時に読み出す。そして、異常データの検出後は、読み出しコントローラ 22 a は、HDD # 0 から奇数の符号が付された単位データを読み出すとともに、HDD # 1 から偶数の符号が付された単位データを同時に読み出すよう制御を変更する。

[0034] これにより、第二実施形態では、データ 2 とデータ 3 を一旦破棄して読み直すことによる 1 セクタ分のデータ読み出し分の遅れは生じるものの、その後、読み出しセクタを変更して HDD # 0 と HDD # 1 から同時に連続する単位データを読み出すことができるため、異常検出時においても通常時と同様に、高速なデータ読み出しを継続することができる。

[0035] (HDD 障害時)

次に、2 台の HDD のうちの一方の HDD (以下の例では、HDD # 0) 全体が固定的に故障した場合の動作について説明する。図 9 は、本発明の第二実施形態に係るデータ入出力方法を説明する説明図である。HDD # 0 が故障している場合、通常の RAID 1 構成と同様に、RAID コントローラ 2 は、HDD # 0 を閉塞する。読み出し要求元から 8 セクタ分のデータ書き込み要求があると、書き込みコントローラ 21 a は、閉塞されていない HDD # 1 にのみ書き込み処理を行う。このとき、上記で説明したように、書き込みコントローラ 21 a は、複数の単位データをセクタ長に応じた並べ替えを行ってセクタ長情報とともに HDD # 1 へ書き込む。例えば、書き込みコントローラ 21 a は、データ 0 乃至データ 7 をデータ 1、データ 3、データ 5、データ 7、データ 0、データ 2、データ 4、データ 6 の順に並び替えて HDD # 1 に書き込む。このように、通常時と同様にデータの並び替えを行って HDD # 1 に書き込むことにより、リビルド処理後に、図 6 や図 7 (a) で例示したディスク構成を再現することができる。

[0036] そして、8 セクタ分のデータの読み出し要求があると、読み出しコントローラ 22 a は、HDD # 1 からのみデータの読み出し処理を行う。ここで、

読み出しコントローラ 22 a は、書き込み前の元の配列順に HDD # 1 から単位データを読み出して、読み出し要求元（ホスト装置 1）へ送信する。例えば、読み出しコントローラ 22 a は、セクタ 4、セクタ 0、セクタ 5、セクタ 1、セクタ 6、セクタ 2、セクタ 7、セクタ 3 の順にデータ 0 乃至データ 7 を読み出して、単位データ読み出し毎に、ホスト装置 1 へ送信する。

[0037] （リビルド処理）

図 10 は、本発明の第二実施形態によるハードディスクのリビルド処理を説明する説明図である。ここで、HDD # 0 が故障した場合に、HDD # 0 を新たな HDD # 0 に交換してリビルド処理を行う。具体的には、上記の「HDD 障害時」に説明したように、読み出しコントローラ 22 a は、元の配列（データ 0 乃至データ 7）を再現するように、HDD # 1 の適切なセクタから複数の単位データを読み出して、セクタ長（この例では、「8」）の情報とともに書き込みコントローラ 21 a へ送る。書き込みコントローラ 21 a は、上述の「データの書き込み処理」で説明したように、セクタ長に応じた HDD # 0 用のデータ並び替えを行って、HDD # 0 へ書き込む。HDD # 0 に対するリビルド処理が終了すると、HDD # 0 と HDD # 1 には図 6 や図 7（a）に例示したような配列で、連続するセクタに単位データが書き込まれる。その後、読み出しコントローラ 22 a は、HDD # 0 と HDD # 1 の同一セクタから単位データを同時に読み出し、セクタ順にデータ読み出しを実行する。これにより、第二実施形態によるハードディスクの冗長性の確保と、高速な読み出し処理を実現することができる。

[0038] 第二実施形態によれば、第一実施形態と同様に、ディスクアレイ装置 3 の可用性と読出し性能の向上を実現することができる。また、第二実施形態では、第一実施形態に比べて、データ読み出し速度を高速化することができる。

[0039] なお、第二実施形態では、ディスクアレイ装置 3 を 2 台の HDD（HDD # 0 と HDD # 1）で構成する例を示したが、HDD の数を 3 台以上に増設しても、同様の書き込み制御と読み出し制御を実現することができる。例え

ば、3台のHDD（HDD#0、HDD#1、HDD#2）に8セクタ長のデータ（データ0乃至データ7）を書き込む場合、図11に示すような配列に単位データを並び替えて3台のHDDに書き込んでもよい。具体的には、書き込みコントローラ21aは、HDD#0について、データ0、データ3、データ6、データ2、データ5、（空き）、データ1、データ4、データ7の順に単位データを並び替えてセクタ0乃至セクタ8に格納する（セクタ5は空きセクタとする）。また、書き込みコントローラ21aは、HDD#1について、データ1、データ4、データ7、データ0、データ3、データ6、データ2、データ5、（空き）の順に単位データを並び替えてセクタ0乃至セクタ8に格納する（セクタ8は空きセクタとする）。さらに、書き込みコントローラ21aは、HDD#2について、データ2、データ5、（空き）、データ1、データ4、データ7、データ0、データ3、データ6の順に単位データを並び替えてセクタ0乃至セクタ8に格納する（セクタ2は空きセクタとする）。

[0040] 図12は、上述の実施形態における制御装置200の最小構成を示すブロック図である。図12に示すように、制御装置200は、少なくとも書込部210と読出部220を備える。書込部210は、複数の単位データにより構成される書き込み対象データを、複数のハードディスクに分散して書き込む。読出部220は、複数のハードディスクから並行して単位データを所定順序で読み出す。例えば、読出部220は、書き込み前の元のデータ配列において、隣接する単位データを異なるハードディスクから同時に読み出す。一例として、読出部220は、複数のハードディスクにおいて1ずつ異なるセクタから単位データを同時に読み出ししてもよい。別の例では、書込部210は、読出部220が複数のハードディスクのそれぞれにおける同一のセクタを始点としてセクタ順に単位データを読み出す場合、書き込み前の元のデータ配列において隣接する単位データを同時に読み出すことができるように単位データの順序を変更して、複数のハードディスクの各々に書き込んでもよい。なお、制御装置200はRAIDコントローラ2に対応し、書込部21

0は書き込みコントローラ21、21aに対応し、読出部220は読み出しコントローラ22、22aに対応する。

[0041] 上述の実施例は、複数のハードディスクドライブ（HDD）をRAID1構成で備えたディスクアレイ装置において、複数の単位データから構成される書き込み対象データを所定のセクタに分散記憶するものであるが、HDD以外のメモリや記録媒体などの記憶媒体を並列に組み合わせたメモリアレイ構成において、複数の単位データを所定の記憶位置に分散記憶するようにしてもよい。また、上述の実施例では、2台のHDDの同一のセクタから一対の連続する単位データを同時に読み出していたが、複数のHDDの第一セクタから第1群の単位データを同時に読み出し、第2群の単位データは、第一セクタに連続する第二セクタから読み出すように、単位データ群ごとに複数のHDDの同一セクタから同時に読み出すようにしてもよい。また、上述の実施形態では、複数の単位データに連続番号を付して、偶数の番号が付されたデータ群と奇数の番号が付されたデータ群に分割して複数のHDDに分散記憶したが、複数の単位データの分割方法は、これに限定されるものではなく、複数の単位データの連続番号を前半部分と後半部分に分割するか、或いは、1つ置きの番号順に分割するなど、所望の分割方法を適用できる。

[0042] 最後に、上述の実施形態において、構成要素を公知の構成要素に適宜置き換えてもよい。また、本発明は上述の実施形態に限定されるものではなく、本発明は、特許請求の範囲に規定される発明の趣旨を逸脱しない範囲における種々の設計変更や改変をも包含するものである。

[0043] 本願は、2018年11月16日に、日本国に出願された特願2018-215857号に基づき優先権を主張し、その内容をここに援用する。

産業上の利用可能性

[0044] 本発明は、複数のハードディスクドライブ（HDD）よりなるミラーリング方式のRAID1構成に適用されるものであるが、RAID1以外のRAID構成にも適用可能である。また、ハードディスクドライブ（HDD）以外の記憶媒体を組み合わせた構成において、複数のデータを異なる記憶位置

(アドレス)に分散配置して、書き込み処理と読み出し処理を行うような情報処理システムにも適用することができる。

符号の説明

- [0045] 1 0 0 情報処理システム
- 1 ホスト装置
 - 2 R A I Dコントローラ
 - 3 ディスクアレイ装置
 - 1 1 オペレーティングシステム (OS)
 - 1 2 ドライバ
 - 1 3 C P U
 - 2 1、2 1 a 書き込みコントローラ
 - 2 2、2 2 a 読み出しコントローラ
 - 2 0 0 制御装置
 - 2 1 0 書込部
 - 2 2 0 読出部

請求の範囲

[請求項1]

複数のハードディスクを備えたディスクアレイシステムに適用される制御装置であって、

複数の単位データからなる書込対象データを、前記複数のハードディスクに分散して書き込む書込部と、

前記複数のハードディスクから並行して前記複数の単位データを同時に読み出す読出部と、を備え、

前記書込部は、前記複数の単位データの配列を変更せずに前記複数のハードディスク内の複数のセクタに分散して書き込み、前記読出部は、前記複数の単位データの配列において連続する単位データ群を前記複数のハードディスクにおける所定のセクタから同時に読み出すとともに、次の単位データ群は前記所定のセクタと異なるセクタから同時に読み出し、

前記複数のハードディスクのうちのハードディスクの一のセクタに記憶された単位データに異常を検出すると、前記単位データの読み出しセクタを他のハードディスクの一のセクタに設定し、他の単位データを読み出すセクタを前記一のハードディスクの他のセクタに設定する、制御装置。

[請求項2]

複数のハードディスクを備えたディスクアレイシステムに適用される制御装置であって、

複数の単位データからなる書込対象データを、前記複数のハードディスクに分散して書き込む書込部と、

前記複数のハードディスクから並行して前記複数の単位データを同時に読み出す読出部と、を備え、

前記書込部は、前記読出部が前記複数のハードディスクの複数のセクタにおいて所定セクタ順序で前記複数の単位データの配列において連続する単位データ群を同時に読み出すことができるように、前記複数の単位データの配列を変更して、前記複数のハードディスクの前記

複数のセクタに書き込み、

前記読出部は、前記複数のハードディスクから前記所定のセクタ順序で前記複数の単位データの配列において連続する単位データ群を同時に読み出し、

前記複数のハードディスクのうち一のハードディスクの一のセクタに記憶された単位データに異常を検出すると、前記単位データの読み出しセクタを他のハードディスクにおける他のセクタに設定し、他の単位データを読み出すセクタを前記一のハードディスクの他のセクタに設定する、制御装置。

[請求項3] 前記複数のハードディスクのうち前記一のハードディスクが故障して、前記一のハードディスクを閉塞して新たなハードディスクに交換した場合、前記書込部は、前記他のハードディスクに書き込まれた前記複数の単位データを同一の順序で前記新たなハードディスクに書き込む、請求項1又は2に記載の制御装置。

[請求項4] 前記書込部は、前記複数の単位データに連続する符号を付している場合、偶数の符号を付した第一データ群と奇数の符号を付した第二データ群に分割して、前記一のハードディスクには前記第一データ群に続いて前記第二データ群を前記複数のセクタに書き込み、前記他のハードディスクには前記第二データ群に続いて前記第一データ群を前記複数のセクタに書き込むようにした、請求項2に記載の制御装置。

[請求項5] 複数のハードディスクを備えたディスクアレイシステムに適用されるデータ入出力方法であって、

複数の単位データからなる書込対象データにおいて、前記複数の単位データの配列を変更せずに前記複数のハードディスク内の複数のセクタに分散して書き込み、

前記複数の単位データの配列において連続する単位データ群を前記複数のハードディスクにおける所定のセクタから同時に読み出すとともに、次の単位データ群は前記所定のセクタと異なるセクタから同時

に読み出し、

前記複数のハードディスクのうち一のハードディスクの一のセクタに記憶された単位データに異常を検出すると、前記単位データの読み出しセクタを他のハードディスクの一のセクタに設定し、他の単位データを読み出すセクタを前記一のハードディスクの他のセクタに設定する、データ入出力方法。

[請求項6]

複数のハードディスクを備えたディスクアレイシステムに適用されるデータ入出力方法であって、

複数の単位データからなる書込対象データにおいて、前記複数のハードディスクの複数のセクタにおいて所定セクタ順序で前記複数の単位データの配列において連続する単位データ群を同時に読み出すことができるように、前記複数の単位データの配列を変更して、前記複数のハードディスクの前記複数のセクタに書き込み、

前記複数のハードディスクから前記所定のセクタ順序で前記複数の単位データの配列において連続する単位データ群を同時に読み出し、

前記複数のハードディスクのうち一のハードディスクの一のセクタに記憶された単位データに異常を検出すると、前記単位データの読み出しセクタを他のハードディスクにおける他のセクタに設定し、他の単位データを読み出すセクタを前記一のハードディスクの他のセクタに設定する、データ入出力方法。

[請求項7]

前記複数の単位データに連続する符号を付している場合、偶数の符号を付した第一データ群と奇数の符号を付した第二データ群に分割して、前記一のハードディスクには前記第一データ群に続いて前記第二データ群を前記複数のセクタに書き込み、前記他のハードディスクには前記第二データ群に続いて前記第一データ群を前記複数のセクタに書き込むようにした、請求項6に記載のデータ入出力方法。

[請求項8]

複数のハードディスクと、

前記複数のハードディスクに対して複数の単位データの読み出しと書

き込みを制御する制御装置と、を備えたディスクアレイシステムであって、

前記制御装置は、前記複数の単位データの配列を変更せずに前記複数のハードディスク内の複数のセクタに分散して書き込む書込部と、

前記複数の単位データの配列において連続する単位データ群を前記複数のハードディスクにおける所定のセクタから同時に読み出すとともに、次の単位データ群は前記所定のセクタと異なるセクタから同時に読み出す読出部と、を備え、

前記読出部は、前記複数のハードディスクのうちのハードディスクの一のセクタに記憶された単位データに異常を検出すると、前記単位データの読み出しセクタを他のハードディスクの一のセクタに設定し、他の単位データを読み出すセクタを前記一のハードディスクの他のセクタに設定するようにした、ディスクアレイシステム。

[請求項9]

複数のハードディスクと、

前記複数のハードディスクに対して複数の単位データの読み出しと書き込みを制御する制御装置と、を備えたディスクアレイシステムであって、

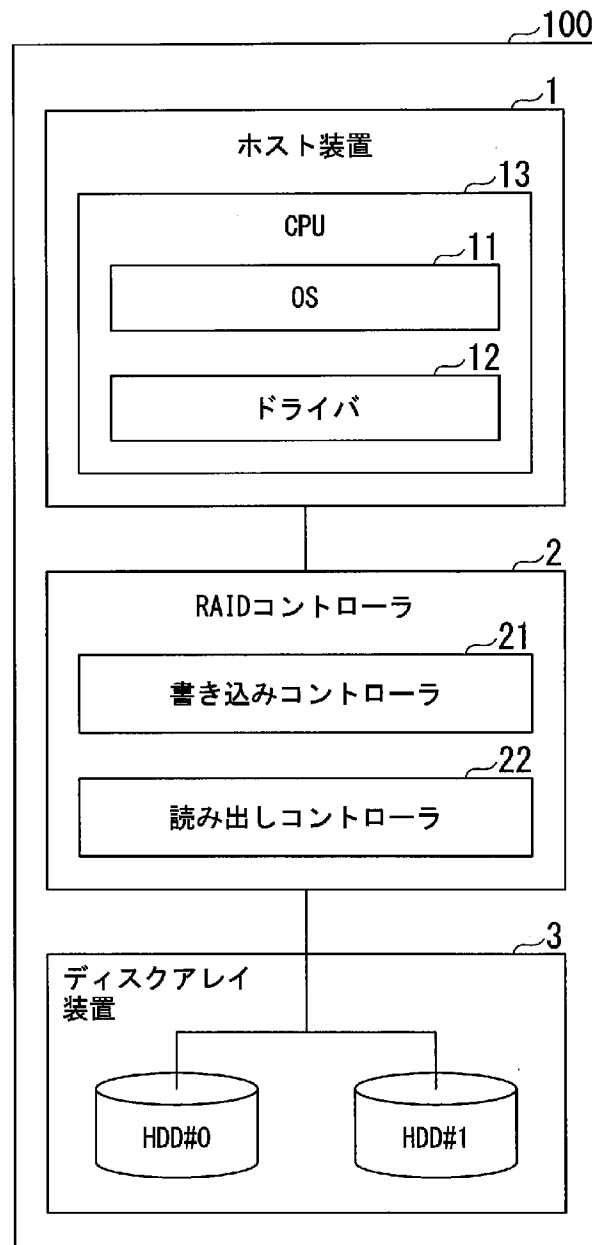
前記制御装置は、前記複数のハードディスクの複数のセクタにおいて所定セクタ順序で前記複数の単位データの配列において連続する単位データ群を同時に読み出すことができるように、前記複数の単位データの配列を変更して、前記複数のハードディスクの前記複数のセクタに書き込む書込部と、

前記複数のハードディスクから前記所定のセクタ順序で前記複数の単位データの配列において連続する単位データ群を同時に読み出す読出部と、を備え、

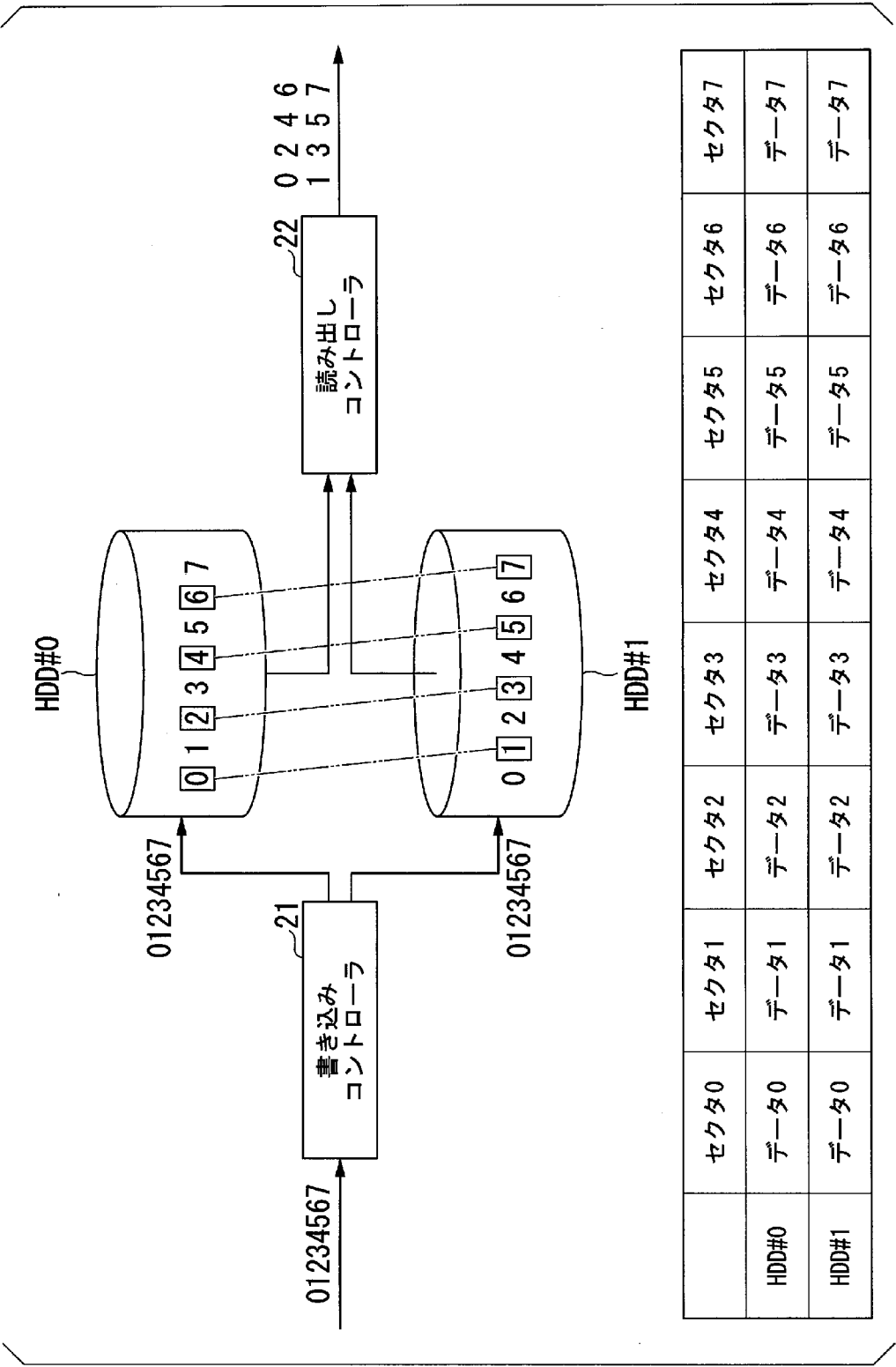
前記読出部は、前記複数のハードディスクのうちのハードディスクの一のセクタに記憶された単位データに異常を検出すると、前記単位データの読み出しセクタを他のハードディスクにおける他のセクタ

に設定し、他の単位データを読み出すセクタを前記一のハードディスクの他のセクタに設定する、ディスクアレイシステム。

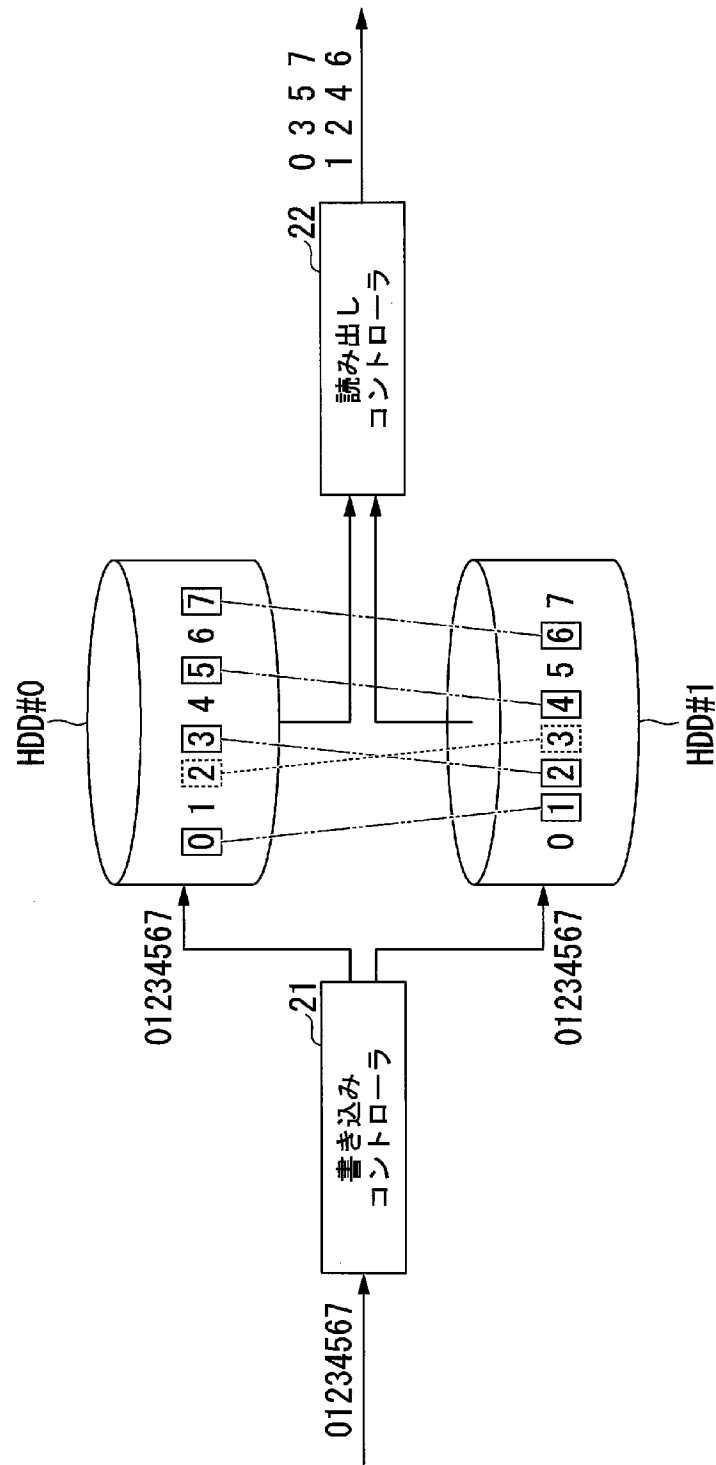
[図1]



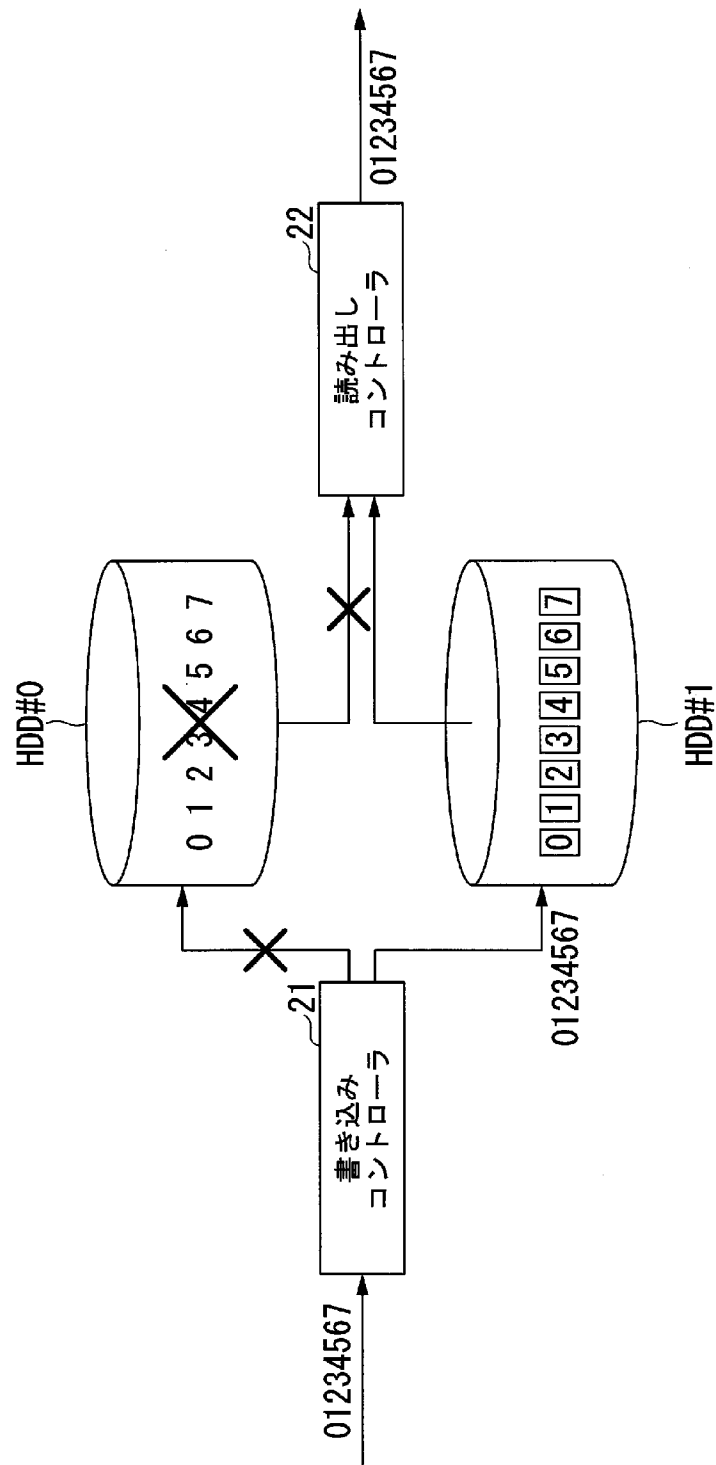
[図2]



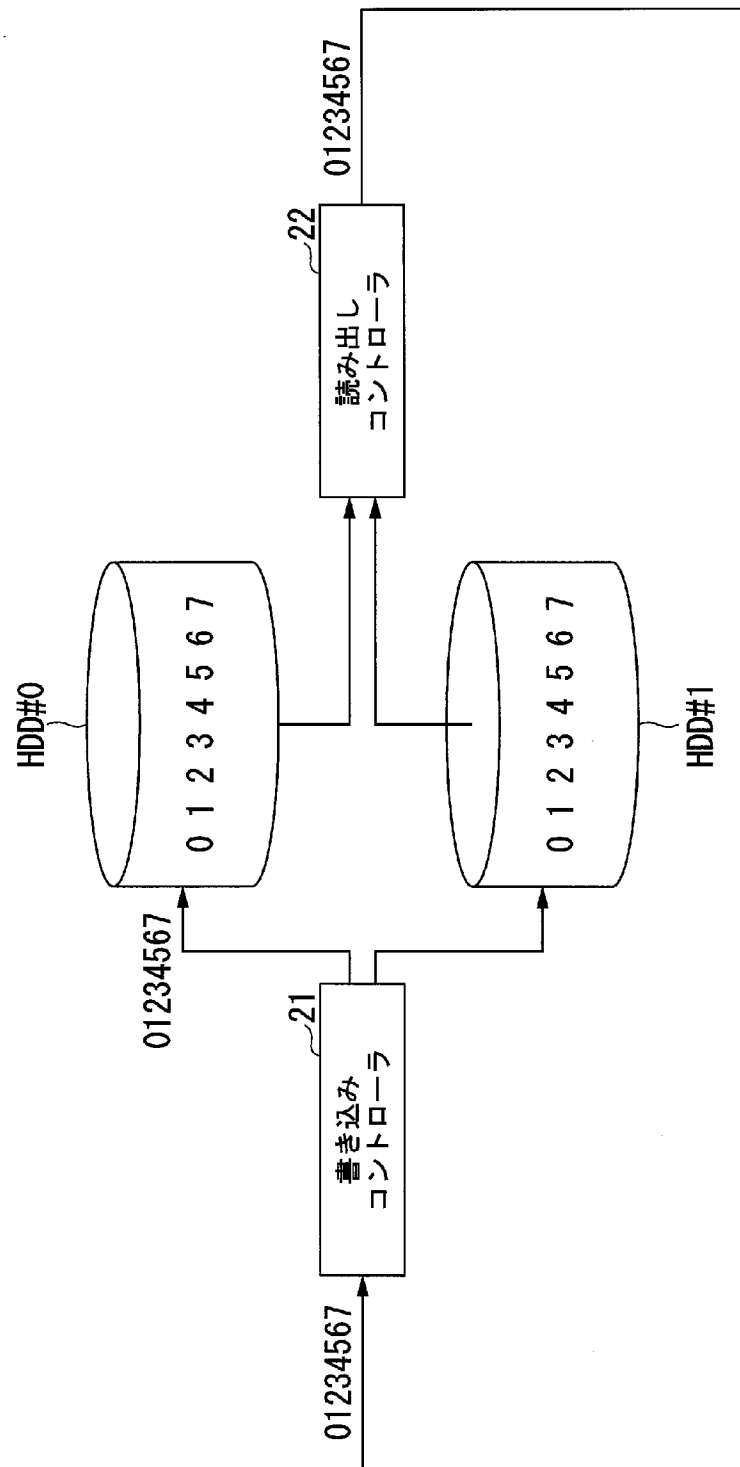
[図3]



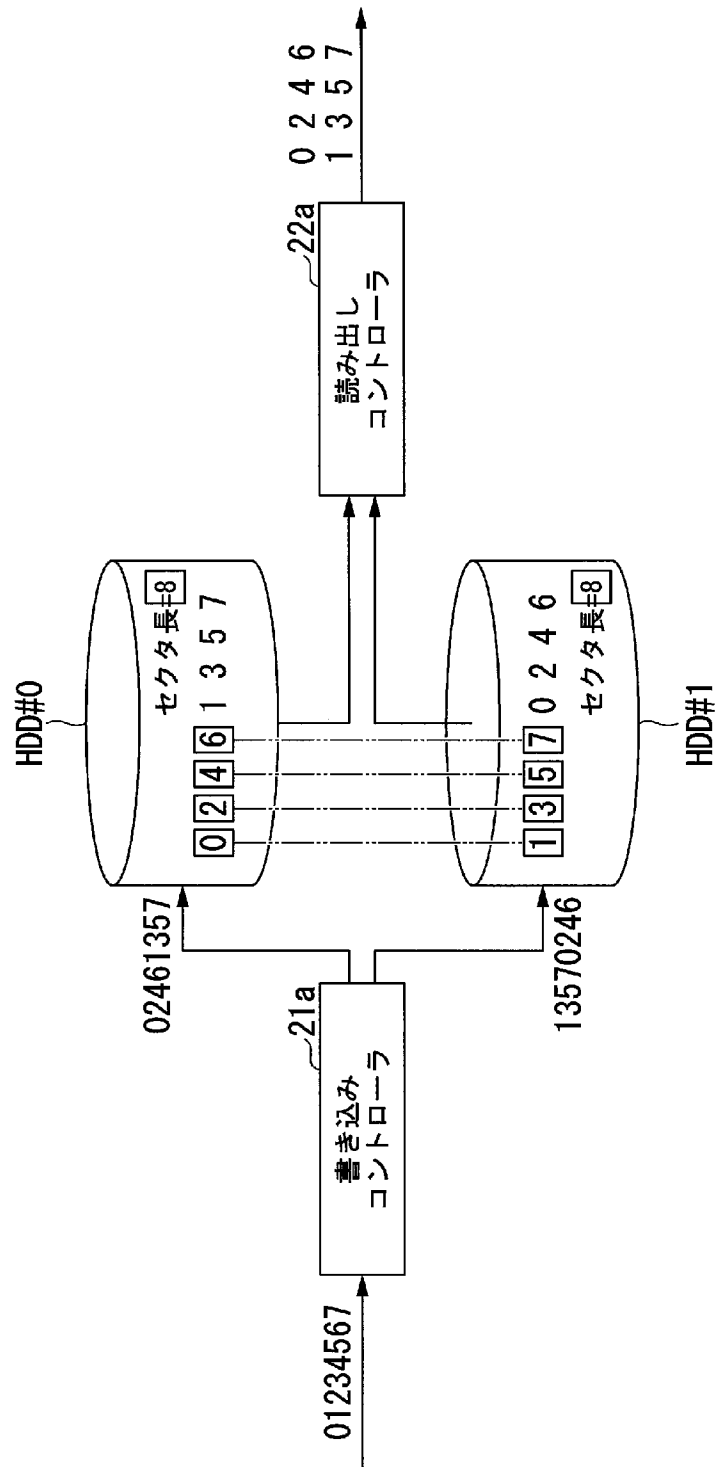
[図4]



[図5]



[図6]



[図7]

データ格納位置(セクタ長=8の場合)

	セクタ0	セクタ1	セクタ2	セクタ3	セクタ4	セクタ5	セクタ6	セクタ7
HDD#0	データ0	データ2	データ4	データ6	データ1	データ3	データ5	データ7
HDD#1	データ1	データ3	データ5	データ7	データ0	データ2	データ4	データ6

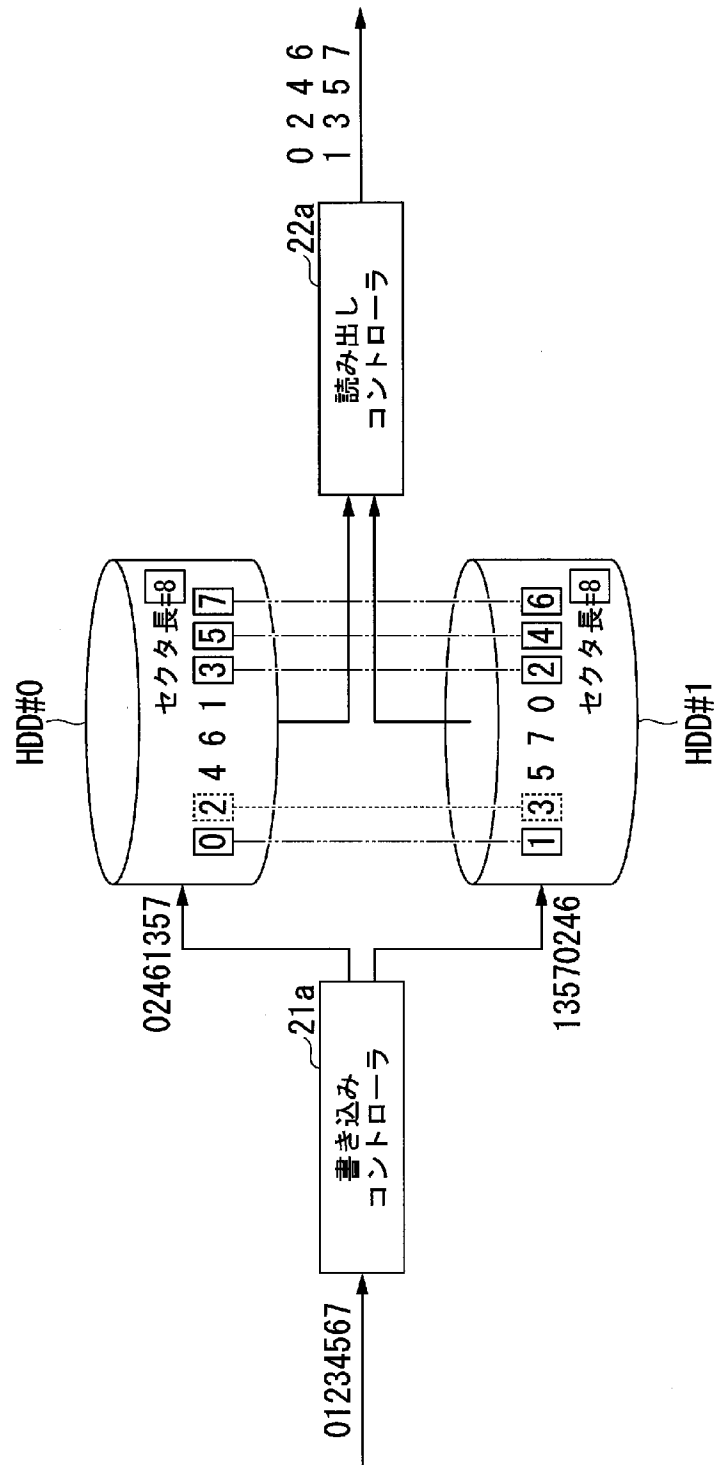
(a)

データ格納位置(セクタ長=4の場合)

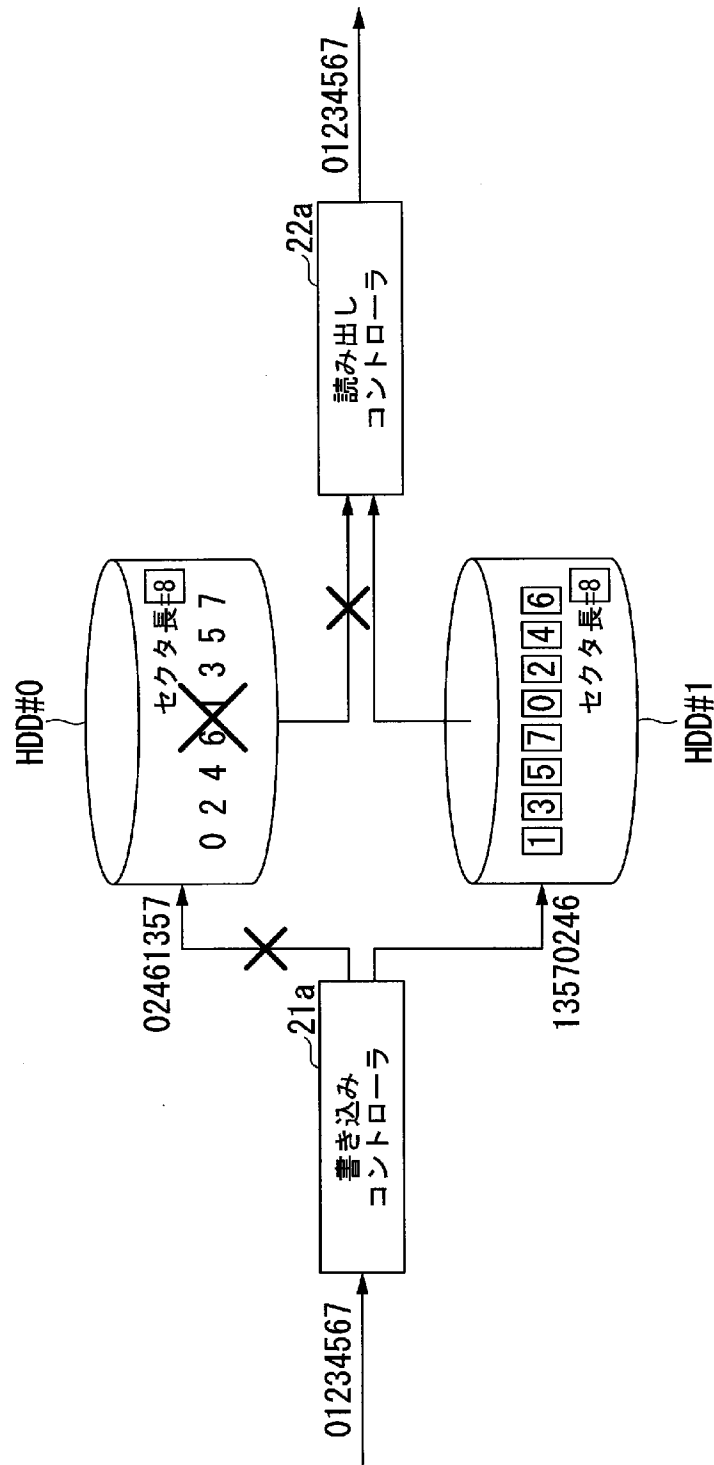
	セクタ0	セクタ1	セクタ2	セクタ3
HDD#0	データ0	データ2	データ1	データ3
HDD#1	データ1	データ3	データ0	データ2

(b)

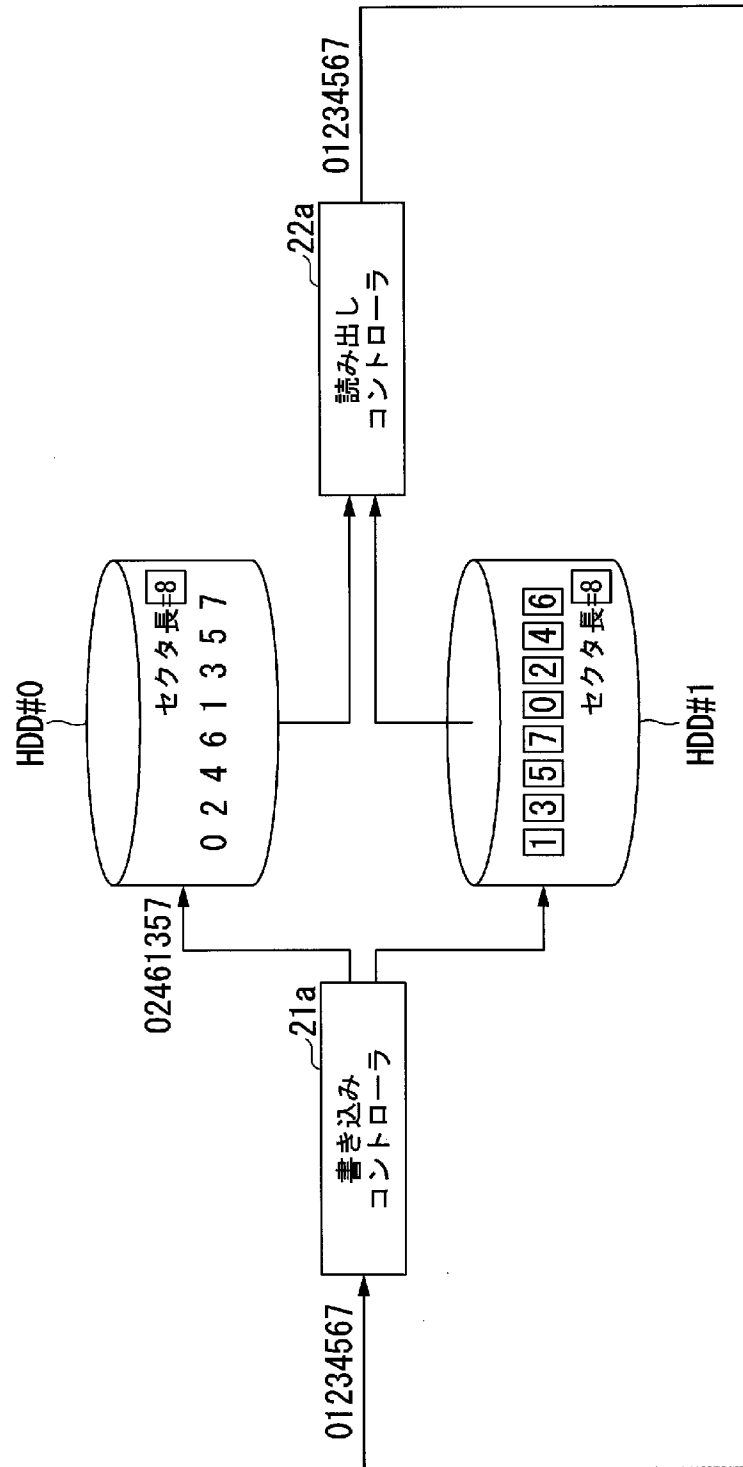
[図8]



[図9]



[図10]

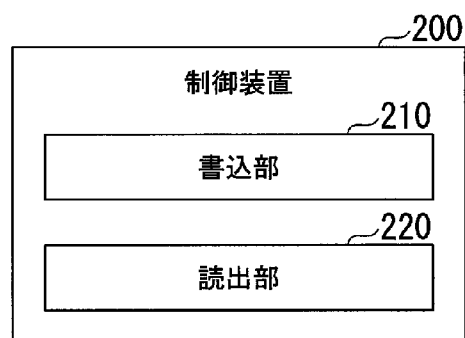


[図11]

HDD3台時のデータ格納位置(セクタ長=8の場合) ×のセクタへは書き込まない

	セクタ0	セクタ1	セクタ2	セクタ3	セクタ4	セクタ5	セクタ6	セクタ7	セクタ8
HDD#0	データ0	データ3	データ6	データ2	データ5	×	データ1	データ4	データ7
HDD#1	データ1	データ4	データ7	データ0	データ3	データ6	データ2	データ5	×
HDD#2	データ2	データ5	×	データ1	データ4	データ7	データ0	データ3	データ6

[図12]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/044304

A. CLASSIFICATION OF SUBJECT MATTER

G06F 11/20 (2006.01) i; G06F 3/06 (2006.01) i

FI: G06F3/06 304E; G06F3/06 302Z; G06F3/06 540; G06F11/20 666

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F11/20; G06F3/06

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2020
Registered utility model specifications of Japan	1996-2020
Published registered utility model applications of Japan	1994-2020

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2003-108316 A (TOSHIBA CORP.) 11.04.2003 (2003-04-11)	1-9
A	JP 2006-507603 A (INTEL CORPORATION) 02.03.2006 (2006-03-02)	1-9
A	US 2009/0216946 A1 (SUN, Zhiming) 27.08.2009 (2009-08-27)	1-9
A	JP 2010-15195 A (TOSHIBA CORP.) 21.01.2010 (2010-01-21)	1-9
A	JP 2012-103856 A (JAPAN BROADCASTING CORPORATION) 31.05.2012 (2012-05-31)	1-9

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
07 January 2020 (07.01.2020)

Date of mailing of the international search report
21 January 2020 (21.01.2020)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/JP2019/044304

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
JP 2003-108316 A	11 Apr. 2003	(Family: none)	
JP 2006-507603 A	02 Mar. 2006	KR 10-2005-0086832 A	
		CN 1503140 A	
US 2009/0216946 A1	27 Aug. 2009	TW 200937190 A	
JP 2010-15195 A	21 Jan. 2010	US 2009/0327803 A1	
JP 2012-103856 A	31 May 2012	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06F 11/20(2006.01)i; G06F 3/06(2006.01)i FI: G06F3/06 304E; G06F3/06 302Z; G06F3/06 540; G06F11/20 666														
B. 調査を行った分野														
調査を行った最小限資料（国際特許分類（IPC）） G06F11/20; G06F3/06														
最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2020年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2020年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2020年</td> </tr> </table>			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2020年	日本国実用新案登録公報	1996 - 2020年	日本国登録実用新案公報	1994 - 2020年				
日本国実用新案公報	1922 - 1996年													
日本国公開実用新案公報	1971 - 2020年													
日本国実用新案登録公報	1996 - 2020年													
日本国登録実用新案公報	1994 - 2020年													
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）														
C. 関連すると認められる文献														
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号												
A	JP 2003-108316 A（株式会社東芝）11.04.2003（2003 - 04 - 11）	1-9												
A	JP 2006-507603 A（インテル コーポレーション）02.03.2006（2006 - 03 - 02）	1-9												
A	US 2009/0216946 A1（SUN, Zhi-Ming）27.08.2009（2009 - 08 - 27）	1-9												
A	JP 2010-15195 A（株式会社東芝）21.01.2010（2010 - 01 - 21）	1-9												
A	JP 2012-103856 A（日本放送協会）31.05.2012（2012 - 05 - 31）	1-9												
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの</td> </tr> <tr> <td>“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの</td> <td>“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの</td> </tr> <tr> <td>“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの</td> <td>“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの</td> </tr> <tr> <td>“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの	“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの	“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの	“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献	“O” 口頭による開示、使用、展示等に言及する文献		“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献	
* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの													
“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの													
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの													
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献													
“O” 口頭による開示、使用、展示等に言及する文献														
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献														
国際調査を完了した日 07.01.2020	国際調査報告の発送日 21.01.2020													
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 吉田 歩 5S 1206 電話番号 03-3581-1101 内線 3546													

国際調査報告
 パテントファミリーに関する情報

国際出願番号

PCT/JP2019/044304

引用文献			公表日	パテントファミリー文献			公表日
JP	2003-108316	A	11.04.2003	(ファミリーなし)			
JP	2006-507603	A	02.03.2006	KR	10-2005-0086832	A	
				CN	1503140	A	
US	2009/0216946	A1	27.08.2009	TW	200937190	A	
JP	2010-15195	A	21.01.2010	US	2009/0327803	A1	
JP	2012-103856	A	31.05.2012	(ファミリーなし)			