

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6089844号
(P6089844)

(45) 発行日 平成29年3月8日(2017.3.8)

(24) 登録日 平成29年2月17日(2017.2.17)

(51) Int.Cl.

F I

G 0 6 F 3 / 0 6 (2006.01)

G 0 6 F 3 / 0 6 3 0 2 E

請求項の数 7 (全 38 頁)

(21) 出願番号 特願2013-59325 (P2013-59325)
 (22) 出願日 平成25年3月22日 (2013. 3. 22)
 (65) 公開番号 特開2014-186412 (P2014-186412A)
 (43) 公開日 平成26年10月2日 (2014. 10. 2)
 審査請求日 平成27年11月6日 (2015. 11. 6)

(73) 特許権者 000005223
 富士通株式会社
 神奈川県川崎市中原区上小田中4丁目1番
 1号
 (74) 代理人 100092978
 弁理士 真田 有
 (74) 代理人 100112678
 弁理士 山本 雅久
 (72) 発明者 小林 秀史
 神奈川県川崎市中原区上小田中4丁目1番
 1号 株式会社富士通コンピュータテクノ
 ロジーズ内

最終頁に続く

(54) 【発明の名称】 制御装置、ストレージ装置、及び制御プログラム

(57) 【特許請求の範囲】

【請求項 1】

入力される第1サイズ単位のデータブロックを、記憶装置の複数のデータ領域のいずれかに対して前記第1サイズと異なる第2サイズ単位で書き込む制御を行なう書込制御部を備え、

前記書込制御部は、

複数のデータブロックを1単位として、当該複数のデータブロックに対応する複数のブロックアドレスを前記記憶装置の前記複数のデータ領域の各々にマッピングするマッピング部と、

各データ領域にマッピングされた前記複数のデータブロックの各々のデータ保存状態を示す管理情報を管理する管理部と、

書き込み対象のデータブロックに対応するブロックアドレスがマッピングされた複数のデータ領域の中から、前記管理情報に基づいて、前記書き込み対象のデータブロックの書き込み先となる一のデータ領域を判定する書込領域判定部と、

前記書き込み対象のデータブロックのデータを含む前記第2サイズの書込用データを生成し、生成した前記書込用データを、前記書込領域判定部により判定された前記一のデータ領域に書き込む処理を行なう書込処理部と、
を備えることを特徴とする、制御装置。

【請求項 2】

前記管理部は、

10

20

前記書き込み対象のデータブロックに対応するブロックアドレスがマッピングされた複数のデータ領域において、前記書き込み対象のデータブロックのデータ保存状態が、前記書き込み対象のデータブロックが書き込まれた前記一のデータ領域では有効を示し、他の全てのデータ領域では無効を示すように、前記管理情報を書き換える
ことを特徴とする、請求項 1 記載の制御装置。

【請求項 3】

前記管理部は、

前記複数のデータ領域の各々にマッピングされた前記複数のブロックアドレスについて、データブロックごとに、前記複数のデータ領域のうちのいずれかのデータ領域内のデータブロックが、他の全てのデータ領域内の当該データブロックとの間で排他的に有効を示すように、前記データ保存状態を管理し、

10

前記書込領域判定部は、

前記書き込み対象のデータブロックに対応するブロックアドレスがマッピングされた複数のデータ領域の前記管理情報に基づいて、前記複数のデータ領域の一つである第 1 のデータ領域における全てのデータブロックのデータ保存状態が無効であると判断した場合に、前記第 1 のデータ領域を、前記一のデータ領域と判定する
ことを特徴とする、請求項 2 記載の制御装置。

【請求項 4】

所定のタイミングで、前記管理部が管理する前記データ保存状態に基づいて、前記複数のデータ領域の各々から前記データ保存状態が有効な前記第 1 サイズのデータブロックを
全て読み出し、読み出した前記データブロックを含む前記第 2 サイズの書込用データを生成して、生成した前記書込用データを、前記複数のデータ領域のいずれかのデータ領域に書き込む最適化部、

20

をさらに備えることを特徴とする、請求項 2 又は請求項 3 記載の制御装置。

【請求項 5】

読み出し対象のデータである一以上のデータブロックに対応するブロックアドレスがマッピングされた複数のデータ領域の前記管理情報に基づいて、前記複数のデータ領域の中から一以上のデータ領域を判定する読出領域判定部と、

前記読み出し対象の一以上のデータブロックに対応するブロックアドレスがマッピングされた複数のデータ領域の前記管理情報に基づいて、前記読出領域判定部により判定された前記一以上のデータ領域の中から、前記データ保存状態が有効な前記読み出し対象の一以上のデータブロックを読み出す処理を行なう読出処理部と、

30

をさらに備えることを特徴とする、請求項 2 ～ 4 のいずれか 1 項記載の制御装置。

【請求項 6】

データを記憶する記憶装置と、

入力される第 1 サイズ単位のデータブロックを、前記記憶装置の複数のデータ領域のいずれかに対して前記第 1 サイズと異なる第 2 サイズ単位で書き込む制御を行なう制御装置と、をそなえ、

前記制御装置は、

複数のデータブロックを 1 単位として、当該複数のデータブロックに対応する複数のブロックアドレスを前記記憶装置の前記複数のデータ領域の各々にマッピングするマッピング部と、

40

各データ領域にマッピングされた前記複数のデータブロックの各々のデータ保存状態を示す管理情報を管理する管理部と、

書き込み対象のデータブロックに対応するブロックアドレスがマッピングされた複数のデータ領域の中から、前記管理情報に基づいて、前記書き込み対象のデータブロックの書き込み先となる一のデータ領域を判定する書込領域判定部と、

前記書き込み対象のデータブロックのデータを含む前記第 2 サイズの書込用データを生成し、生成した前記書込用データを、前記書込領域判定部により判定された前記一のデータ領域に書き込む処理を行なう書込処理部と、

50

を備えることを特徴とする、ストレージ装置。

【請求項 7】

入力される第 1 サイズ単位のデータブロックを、記憶装置の複数のデータ領域のいずれかに対して前記第 1 サイズと異なる第 2 サイズ単位で書き込む制御を行なうコンピュータに、

複数のデータブロックを 1 単位として、当該複数のデータブロックに対応する複数のブロックアドレスを前記記憶装置の前記複数のデータ領域の各々にマッピングし、

各データ領域にマッピングされた前記複数のデータブロックの各々のデータ保存状態を示す管理情報を管理し、

書き込み対象のデータブロックに対応するブロックアドレスがマッピングされた複数のデータ領域の中から、前記管理情報に基づいて、前記書き込み対象のデータブロックの書き込み先となる一のデータ領域を判定し、

前記書き込み対象のデータブロックのデータを含む前記第 2 サイズの書込用データを生成し、生成した前記書込用データを、判定した前記一のデータ領域に書き込む、
処理を実行させることを特徴とする、制御プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、制御装置、ストレージ装置、及び制御プログラムに関する。

【背景技術】

【0002】

従来、512 エミュレーションを行なう 4 K セクタ Hard Disk Drive (HDD) が知られている。4 K セクタ HDD は、ディスク (媒体) に記録されるデータの最小単位である 1 セクタが 4 K バイト (4096 バイト (4 KiB)) の HDD である。

512 エミュレーションは、ディスク上では 1 セクタ = 4096 バイトであるが、PC やサーバ等のホスト装置に対する I/F (インタフェース) 上では、512 バイト / セクタとなるように行なわれるエミュレーションである。以下、512 エミュレーションを行なう 4 K セクタ HDD を 4 K セクタ HDD (512 E) という。

【0003】

4 K セクタ HDD が登場する以前から、1 セクタが 512 バイトである HDD が用いられており、ホスト装置には、512 バイト / セクタの HDD に合わせて 512 バイト単位で HDD に対して書込コマンドを発行するものが少なくない。

4 K セクタ HDD (512 E) は、上記のような 512 バイト単位で書込コマンドを発行するホスト装置との互換性を担保するために、ホスト装置に対して 512 バイト / セクタの HDD と同様な I/F を提供する。従って、ホスト装置は、512 バイト / セクタの HDD へのアクセスと同様のコマンド及びデータサイズで、4 K セクタ HDD (512 E) へのアクセスを行なうことができる。

【0004】

ここで、4 K セクタ HDD (512 E) では、ホスト装置からの書込コマンドに応じて、Read Modify Write (リードモディファイライト; 以下、RMW という) が発生することがある。以下、図 14 及び図 15 を参照して、RMW について説明する。

図 14 は、4 K セクタ HDD (512 E) による RMW の動作例を示す図であり、図 15 は、4 K セクタ HDD (512 E) における Logical Block Address (LBA) とセクタとの対応関係を例示する図である。

【0005】

図 14 に示すように、情報処理システム 100 は、ストレージ装置 200 及びホスト装置 300 を備える。ホスト装置 300 は、PC やサーバ等の情報処理装置であり、512 バイト単位で書込コマンド及び読出コマンドを発行する。ストレージ装置 200 は、コントローラモジュール 400 及び 4 K セクタ HDD (512 E) 500 を備える。以下、4 K セクタ HDD (512 E) 500 を単に HDD 500 ともいう。

【 0 0 0 6 】

コントローラモジュール 4 0 0 は、ホスト装置 3 0 0 からのアクセス要求に応じて H D D 5 0 0 に対してアクセスを行なう制御装置であり、H D D 5 0 0 に出力する書込データ及び H D D 5 0 0 から読み出した読出データ等を一時的に格納するメモリ 4 1 0 を備える。H D D 5 0 0 は、4 K セクタ H D D (5 1 2 E) であり、コントローラモジュール 4 0 0 との間でやり取りされる書込 / 読出データを一時的に格納するメモリ 5 1 0、及び、データを記憶する記憶領域 (ディスク) 5 2 0 を備える。

【 0 0 0 7 】

H D D 5 0 0 は、図 1 5 に示すように、ホスト装置 3 0 0 に対する I / F 上では 5 1 2 バイトを 1 ブロックとして管理する一方、ディスク上では 4 0 9 6 バイトを 1 ブロック (セクタ) として管理する。ホスト装置 3 0 0 は、コントローラモジュール 4 0 0 配下に接続された H D D 5 0 0 へアクセスを行なう際には、H D D 5 0 0 の I / F によりホスト装置 3 0 0 が認識できる 5 1 2 バイト単位の L B A を指定して、コマンドを発行する。

10

【 0 0 0 8 】

なお、図 1 5 に示す例では、L B A の # 0 0 ~ # 1 5 がそれぞれ 5 1 2 バイトのブロックであり、セクタ # 0 及び # 1 がそれぞれ 4 0 9 6 バイトのブロックである。また、L B A の # 0 0 ~ # 0 7 と H D D 5 0 0 のセクタ # 0 とが対応し、L B A の # 0 8 ~ # 1 5 と H D D 5 0 0 のセクタ # 1 とが対応する。

図 1 4 に示す情報処理システム 1 0 0 において、ホスト装置 3 0 0 から 5 1 2 バイトの書込データを L B A # 0 2 に書き込む書込コマンドが発行されると、コントローラモジュール 4 0 0 は、書込データを一旦メモリ 4 1 0 に溜め込む。このとき、コントローラモジュール 4 0 0 は、書込データを 5 1 2 バイト単位のデータブロックで管理する。

20

【 0 0 0 9 】

そして、コントローラモジュール 4 0 0 は、メモリ 4 1 0 に溜め込んだ書込データを非同期で H D D 5 0 0 へ書き込む (図 1 4 中の矢印 (1) 参照)。H D D 5 0 0 は、書込データを受信すると、書き込み対象のブロックである L B A # 0 2 を含む 4 0 9 6 バイトのデータブロック # 0 を、記憶領域 5 2 0 からメモリ 5 1 0 に読み出す (図 1 4 中の矢印 (2) 参照)。

【 0 0 1 0 】

次いで、H D D 5 0 0 は、記憶領域 5 2 0 から読み出した 4 0 9 6 バイトのデータとコントローラモジュール 4 0 0 から受信した書込データとをマージする (図 1 4 中の (3) 参照)。つまり、H D D 5 0 0 は、読み出した 4 0 9 6 バイトのデータのうち、L B A # 0 2 に対応する領域の 5 1 2 バイトのデータを、書込データで上書き (モディファイ) する。

30

【 0 0 1 1 】

最後に、H D D 5 0 0 は、メモリ 5 1 0 に保持されたマージ後の 4 0 9 6 バイトのデータを記憶装置 5 2 0 へ書き込む (図 1 4 中の矢印 (4) 参照)。

以上のように、4 K セクタ H D D (5 1 2 E) は、1 セクタである 4 0 9 6 バイト未満の書込データを受信すると、図 1 4 中の (2) ~ (4) に示す R M W を行なう。

なお、関連する技術として、未記録 Error Correction Code (E C C) ブロックを含む光ディスク等に、リードモディファイを行わずにユーザデータを記録可能にする技術が知られている (例えば、特許文献 1 参照)。この技術では、光ディスク装置は、未使用 E C C ブロック内のセクタにユーザデータを記録するコマンドを受け取った場合、R M W を行わずにホスト装置から転送されたユーザデータにダミーデータを付加する。そして、光ディスク装置は、ダミーデータの付加によりユーザデータを 1 E C C ブロック分のデータ長とした上で目的の E C C ブロックに記録する。

40

【 0 0 1 2 】

また、関連する他の技術として、4 K バイトでディスクに読み書きするハードディスクドライブ装置において、ディスクに記憶されている 5 1 2 バイトの管理データを上書きする際の処理速度を向上させる技術が知られている (例えば、特許文献 2 参照)。この技術

50

では、ハードディスクドライブ装置は、管理データを含む4Kバイトのデータを予めキャッシュメモリに記憶しておき、データの書込命令により、PCからの管理データでキャッシュメモリ内の管理データを更新する。そして、ハードディスクドライブ装置は、更新された管理データを含む4Kバイトのデータをキャッシュメモリからディスクに書き込む。

【先行技術文献】

【特許文献】

【0013】

【特許文献1】特開平10-106170号公報

【特許文献2】特開2012-113789号公報

【発明の概要】

10

【発明が解決しようとする課題】

【0014】

4KセクタHDD(512E)500において上述したRMWが発生すると、512バイト/セクタのHDDと比べて、HDDの性能低下が懸念される。

具体的には、512バイト/セクタのHDDは、例えば512バイトの書込データを受け取ると、書き込みのためのシーク処理、及び、書込データの書込処理を行なう。これに対し、4KセクタHDD(512E)は、RMWが発生すると、512バイト/セクタのHDDにおける処理に加えて、読み出しのためのシーク処理、データの読出処理、及び、マージ処理を行なう(図14の(2)~(4)参照)。このように、RMWの処理は、512バイト/セクタのHDDによるデータの書き込みよりもオーバーヘッドが大きい。

20

【0015】

特に、数ms程度の時間がかかるシーク処理に着目すると、512バイト/セクタのHDDでは書き込みのために1回行なわれるのに対し、RMWにおいては読み出し及び書き込みのために2回行なわれることになる。

以上のように、4KセクタHDD(512E)によるRMWは、HDD単体でみた場合、512バイト/セクタのHDDによるデータの書き込みよりもHDD性能を低下させる要因となる。

【0016】

また、上述したユーザデータにダミーデータを付加する技術は、ダミーデータが付加されたユーザデータを、未使用の1ECCブロック内へ記録するものであり、既にデータが記憶されているブロック内への記録については、考慮されていない。

30

さらに、上述した4Kバイトのデータを予めキャッシュメモリに記憶しておく技術では、キャッシュメモリの占有時間及び占有率が高くなり、ディスクに対するキャッシュとして使用可能なキャッシュメモリの領域が減少する。キャッシュメモリの利用可能領域が減少すると、同容量のキャッシュメモリを有するHDDよりもキャッシュミスの可能性が高くなるため、HDDの性能低下が懸念される。

【0017】

ここまで、4KセクタHDD(512E)においてRMWが発生する場合を例に挙げて説明したが、入力される最小サイズ(単位)と異なるサイズで書込データを保存し、RMWに類する処理を行なう他の記憶装置においても同様の問題がある。このような記憶装置の一例としては、4Kバイト等のページ単位で書込データを記憶するSolid State Drive(SSD)やUniversal Serial Bus(USB)メモリ等のフラッシュメモリが挙げられる。

40

【0018】

1つの側面では、本発明は、第1サイズの単位データを第1サイズと異なる第2サイズのデータ領域に保存する記憶装置における書込処理の性能低下を軽減させることを目的とする。

なお、前記目的に限らず、後述する発明を実施するための形態に示す各構成により導かれる作用効果であって、従来の技術によっては得られない作用効果を奏することも本発明の他の目的の1つとして位置付けることができる。

50

【課題を解決するための手段】

【0019】

本件の制御装置は、入力される第1サイズ単位のデータブロックを、記憶装置の複数のデータ領域のいずれかに対して前記第1サイズと異なる第2サイズ単位で書き込む制御を行なう書込制御部を備え、前記書込制御部は、複数のデータブロックを1単位として、当該複数のデータブロックに対応する複数のブロックアドレスを前記記憶装置の前記複数のデータ領域の各々にマッピングするマッピング部と、各データ領域にマッピングされた前記複数のデータブロックの各々のデータ保存状態を示す管理情報を管理する管理部と、書き込み対象のデータブロックに対応するブロックアドレスがマッピングされた複数のデータ領域の中から、前記管理情報に基づいて、前記書き込み対象のデータブロックの書き込み先となる一のデータ領域を判定する書込領域判定部と、前記書き込み対象のデータブロックのデータを含む前記第2サイズの書込用データを生成し、生成した前記書込用データを、前記書込領域判定部により判定された前記一のデータ領域に書き込む処理を行なう書込処理部と、を備える。

10

【発明の効果】

【0020】

一実施形態によれば、第1サイズの単位データを第1サイズと異なる第2サイズのデータ領域に保存する記憶装置における書込処理の性能低下を軽減させることができる。

【図面の簡単な説明】

【0021】

20

【図1】一実施形態に係る通信システムの構成例を示す図である。

【図2】図1に示すHDDにおけるLBAとセクタとの対応関係を例示する図である。

【図3】図1に示す処理部による動作例を説明するための図である。

【図4】図1に示す処理部による動作例を説明するための図である。

【図5】図1に示す処理部による動作例を説明するための図である。

【図6】図1に示す処理部による動作例を説明するための図である。

【図7】図1に示す処理部による動作例を説明するための図である。

【図8】図1に示す処理部による動作例を説明するための図である。

【図9】(a)及び(b)は、それぞれ、図7及び図8に示す動作例におけるHDDでのマージ量を説明するための図である。

30

【図10】図1に示す書込制御部による動作例を説明するフローチャートである。

【図11】図1に示す読出制御部による動作例を説明するフローチャートである。

【図12】一実施形態に係る情報処理システムの構成の変形例を示す図である。

【図13】図1又は図12に示すコントローラモジュールのハードウェア構成例を示す図である。

【図14】HDDによるリードモディファイライトの動作例を示す図である。

【図15】4KセクタHDD(512エミュレーション)におけるLBAとセクタとの対応関係を例示する図である。

【発明を実施するための形態】

【0022】

40

以下、図面を参照して実施の形態を説明する。

〔1〕一実施形態

〔1-1〕情報処理システムの説明

図1は、一実施形態に係る情報処理システム1の構成例を示す図である。

図1に示すように、情報処理システム1は、ストレージ装置2及びホスト装置3を備える。ホスト装置3は、PCやサーバ等の情報処理装置であり、ストレージ装置2に対して512バイト単位で書込又は読出コマンドを発行する上位装置の一例である。ストレージ装置2は、コントローラモジュール4、少なくとも一つの4KセクタHDD(512E)5、チャネルアダプタ6、及び、デバイスアダプタ7を備える。以下、4KセクタHDD(512E)5を単にHDD5ともいう。

50

【 0 0 2 3 】

チャネルアダプタ 6 は、ホスト装置 3 と接続され、ホスト装置 3 とのインタフェース制御を行なうモジュールであり、ホスト装置 3 との間で通信やデータ転送を行なう。デバイスアダプタ 7 は、ストレージ装置 2 に収容される H D D 5 とのインタフェース制御を行なうモジュールであり、H D D 5 との間で通信やデータ転送を行なう。

コントローラモジュール（制御装置，制御部）4 は、ホスト装置 3 からのアクセス要求に応じて H D D 5 に対してアクセスを行なう制御装置であり、後述するメモリ 2 0 や各種資源の管理及びライトバック処理等の制御を行なうモジュールである。本実施形態に係るコントローラモジュール 4 は、ホスト装置 3 に対して第 1 サイズのデータブロック単位で入出力を行ない、H D D 5 の各データ領域に対して第 1 サイズと異なる第 2 サイズ単位でデータの入出力を行なう。

10

【 0 0 2 4 】

H D D（記憶装置）5 は、図示しない Central Processing Unit（C P U）等のプロセッサを備え、上述した 5 1 2 エミュレーションを行なう 4 K セクタ H D D である。また、H D D 5 は、メモリ 5 1、記憶領域 5 2、及び、記憶装置ビットマップ 5 3 を備える。なお、H D D 5 に代えて、4 K バイト等のページ単位で書込データを記憶し、R M W に類する処理を行なう S S D や U S B メモリ等のフラッシュメモリが用いられてもよい。また、図 1 では、一つの H D D 5 のみを図示しているが、複数の H D D 5 がデバイスアダプタ 7 を介してコントローラモジュール 4 と接続されてもよい。

【 0 0 2 5 】

20

メモリ 5 1 は、コントローラモジュール 4 との間でやり取りされる書込 / 読出データを一時的に格納するものであり、例えば、Random Access Memory（R A M）等の揮発性メモリ又はフラッシュメモリ等の不揮発性メモリが挙げられる。なお、メモリ 5 1 は、上述した揮発性又は不揮発性メモリのほか、H D D や S S D 等の記憶装置であってもよく、情報を記憶可能な種々のデバイスが用いられてよい。例えばメモリ 5 1 として、記憶領域 5 2 と同じ媒体上の別領域が用いられてもよい。

【 0 0 2 6 】

記憶領域 5 2 は、種々のデータやプログラムを記憶するディスク領域である。本実施形態に係る記憶領域 5 2 においては、一つの L B A に対して、複数（例えば二つ）のデータ領域が割り当てられる。以下、記憶領域 5 2 について、図 2 を参照して説明する。

30

図 2 は、図 1 に示す H D D 5 における L B A とセクタとの対応関係を例示する図である。H D D 5 は、図 2 に示すように、ホスト装置 3 に対する I / F 上では 5 1 2 バイト（第 1 サイズ）を 1 ホストブロックとして管理する。ホストブロックは、H D D 5 に対するホスト装置 3 からの Input/Output（I / O）命令単位データブロックである。

【 0 0 2 7 】

一方、H D D 5 は、記憶領域 5 2 上では第 1 サイズの整数倍である 4 0 9 6 バイト（第 2 サイズ）を 1 セクタブロックとして記憶領域 5 2 上の各セクタに対応付けて管理する。セクタブロックは、ハードディスクの記憶領域 5 2 に対する H D D 5 のコントローラからの I / O 命令単位データブロックである。なお、図 2 に示すように、L B A の # 0 0 ~ # 1 5 がそれぞれ 5 1 2 バイトのブロックであり、セクタ # 0 - 0 及び # 0 - 1 がそれぞれ 4 0 9 6 バイトのブロックである。また、以下、セクタブロックを単にセクタという。

40

【 0 0 2 8 】

本実施形態に係る H D D 5 では、L B A の # 0 0 ~ # 0 7 と、H D D 5 のセクタ # 0 - 0 及びセクタ # 0 - 1 とが対応する。すなわち、ホスト装置 3 が認識できる 5 1 2 バイト単位の L B A に対して、複数（図 2 に示す例では二つ）の 4 0 9 6 バイトのセクタ（データ領域）が割り当てられるのである。なお、図 2 において図示を省略しているが、L B A の # 0 8 ~ # 1 5 に対しても、H D D 5 の図示しない二つのセクタが割り当てられる。このような割り当ては、後述する割当処理部 1 4 により行なわれる。

【 0 0 2 9 】

以下、記憶領域 5 2 についてより具体的に説明する。

50

上述のように、コントローラモジュール4又はHDD5は、一つのLBA（以下、ホストブロックともいう）を512バイトで管理する。そして、コントローラモジュール4（割当管理部14）又はHDD5は、8ブロックを1単位として記憶領域52の1セクタにマッピングし、さらに同一ホストブロックに対してもう1セクタをマッピングする。以下、同一の8ホストブロックに対応する二つのセクタ（図2に示すセクタ#0-0及び#0-1参照）を、それぞれA領域52b-1及びB領域52b-2という。なお、以下、A領域52b-1及びB領域52b-2を区別しない場合には、データ領域52bという。これらのA領域52b-1及びB領域52b-2は、記憶領域52における互いに異なる領域であり、且つ、いずれのデータ領域52bも第1サイズの複数のデータブロックの各々のLBAに対応する領域を含む。

10

【0030】

記憶領域52において、A領域52b-1及びB領域52b-2は、例えば数十セクタ～数百セクタの単位で、複数のA領域52b-1、複数のB領域52b-2、複数のA領域52b-1、...という形で交互に配置される（図1参照）。つまり、A領域52b-1及びB領域52b-2は、記憶領域52において1セクタごとに交互に配置されるのではなく、連続するホストブロックの書き込み又は読み出しを考慮して、ある程度意味のある単位で複数セクタがまとめられることが好ましい。なお、HDD5のディスクは、トラック当たりのセクタ数や角速度の関係から、内周部よりも外周部の方が性能が良いため、性能差を考慮して、一对のA領域52b-1及びB領域52b-2が離れすぎないようにすることが好ましい。

20

【0031】

A領域52b-1及びB領域52b-2をそれぞれ所定の単位でまとめることにより、ホスト装置3によりシーケンシャルなアクセスが行なわれた場合でも、HDD5の性能低下を抑止することができる。なお、以下、記憶領域52における、A領域52b-1のグループを第1グループ52a-1といい、B領域52b-2のグループを第2グループ52a-2という。

【0032】

ここで、同一の8ホストブロックに対応するA領域52b-1及びB領域52b-2が格納するデータは、ホストブロックごとに排他的な関係を有する。図2に示すLBA#00～#07に着目して説明すると、例えば、LBA#00～#03の最新のデータがA領域52b-1であるセクタ#0-0に格納されている場合、B領域52b-2であるセクタ#0-1に格納されているデータは無効なデータとなる。一方、LBA#04～#07の最新のデータがセクタ#0-1に格納されている場合には、セクタ#0-0に格納されているデータは無効なデータとなる。

30

【0033】

HDD5は、各データ領域の上述のような排他関係に基づく後述するコントローラモジュール4による制御を受けて、第1サイズのデータブロックを第1サイズよりも大きい第2サイズの複数のデータ領域のいずれか一つに保存する。HDD5におけるデータの保存の態様の詳細は、後述する。

記憶装置ビットマップ格納領域53は、A領域52b-1及びB領域52b-2における排他的なデータ保存状態を管理する記憶装置ビットマップ22aを格納する。

40

【0034】

記憶装置ビットマップ（管理情報、第1管理情報）22aは、各データ領域52bにおける複数の第1サイズのデータブロックの記憶領域のデータ保存状態を示す情報である。記憶装置ビットマップ22aは、以下に例示するように、A領域52b-1及び52b-2のいずれか一方のデータ領域52bに含まれるデータブロックごとに該データブロックのデータ保存状態が有効か否かを示すビットマップ情報とすることができる。例えば、LBA#00～#03のデータがA領域52b-1に、LBA#04～#07のデータがB領域52b-2に、それぞれ格納されている場合、LBA#00～#07に関する記憶装置ビットマップ22aが示すデータ保存状態は、以下のようになる。

50

【 0 0 3 5 】

A 領域 5 2 b - 1 : 1 1 1 1 0 0 0 0

B 領域 5 2 b - 2 : 0 0 0 0 1 1 1 1

なお、上記の記憶装置ビットマップ 2 2 a は、1 ビットが 5 1 2 バイトの 1 ホストブロックに相当し、左から L B A の昇順で配置される。また、記憶装置ビットマップ 2 2 a において、“ 1 ” は有効つまり当該セクタにおいてホストブロックのデータが最新であることを示し、“ 0 ” は無効つまり当該セクタに格納されたホストブロックのデータが無効であることを示す。

【 0 0 3 6 】

また、上記では A 領域 5 2 b - 1 及び B 領域 5 2 b - 2 の双方について、L B A # 0 0 ~ # 0 7 に関する記憶装置ビットマップ 2 2 a の例を示したが、それぞれのビットマップは排他関係を有するため、記憶装置ビットマップ 2 2 a には、A 領域 5 2 b - 1 及び B 領域 5 2 b - 2 のいずれか一方のビットマップが設定されればよい。以下、記憶装置ビットマップ 2 2 a には、A 領域 5 2 b - 1 のビットマップが設定されるものとするが、便宜上、記憶装置ビットマップ 2 2 a に B 領域 5 2 b - 2 のビットマップも設定されるものとして説明する場合がある。

【 0 0 3 7 】

記憶装置ビットマップ格納領域 5 3 は、記憶領域 5 2 における第 1 グループ 5 2 a - 1 (全ての A 領域 5 2 b - 1) についての記憶装置ビットマップ 2 2 a を格納する。例えば、H D D 5 における記憶領域 5 2 の容量が 3 0 0 G バイトである場合、記憶装置ビットマップ格納領域 5 3 は、第 1 グループ 5 2 a - 1 についての 3 6 M バイト程度の記憶装置ビットマップ 2 2 a を格納する。

【 0 0 3 8 】

なお、記憶装置ビットマップ格納領域 5 3 は、H D D 5 のディスク領域の一部の領域とすることができる。

〔 1 - 2 〕コントローラモジュールの説明

次に、図 1 を参照しながら、コントローラモジュール 4 の詳細を説明する。

図 1 に示すように、コントローラモジュール 4 は、C P U 1 0、メモリ 2 0、及び、S S D 3 0 を備える。

【 0 0 3 9 】

C P U (処理部) 1 0 は、メモリ 2 0、S S D 3 0、チャンネルアダプタ 6、及び、デバイスアダプタ 7 とそれぞれ接続され、種々の制御や演算を行なう処理装置 (プロセッサ) の一例である。C P U 1 0 は、メモリ 2 0、S S D 3 0、H D D 5、又は、図示しない Read Only Memory (R O M) 等に格納されたプログラムを実行することにより、コントローラモジュール 4 における種々の機能を実現する。本実施形態に係る C P U 1 0 は、メモリ 2 0 等に格納された制御プログラムを実行することにより、後述する書込制御部 1 1、読出制御部 1 2、記憶装置ビットマップ管理部 1 3、割当処理部 1 4、及び、メモリビットマップ管理部 1 5 を備える処理部としての機能を実現する。

【 0 0 4 0 】

メモリ (保持部) 2 0 は、種々のデータやプログラムを一時的に格納するキャッシュメモリ等の記憶装置であって、C P U 1 0 がプログラムを実行する際に、データやプログラムを一時的に格納・展開して用いる。例えば、メモリ 2 0 は、ユーザデータ領域 2 1 及びビットマップ領域 2 2 を含む。なお、メモリ 2 0 としては、例えば R A M 等の揮発性メモリが挙げられる。

【 0 0 4 1 】

ユーザデータ領域 (保持領域) 2 1 は、C P U 2 0 が処理部として機能するための制御プログラム、ホスト装置 3 から H D D 5 へ書き込まれる書込データ、又は、H D D 5 からホスト装置 3 へ読み出される読出データ等を一時的に格納する領域である。ユーザデータ領域 2 1 は、第 2 サイズである 4 0 9 6 バイト以上の領域である。なお、書込データは、A 領域 5 2 b - 1 及び B 領域 5 2 b - 2 のいずれか一つに書き込まれるデータブロック (

10

20

30

40

50

書き込み対象のデータブロック)であり、ホスト装置3から第1サイズである512バイト単位で入力される。

【0042】

ビットマップ領域22は、CPU20により参照される記憶装置ビットマップ22a及びメモリビットマップ22bを保持する領域である。なお、ビットマップ領域22が保持する記憶装置ビットマップ22aは、HDD5の記憶装置ビットマップ格納領域53又はSSD30の記憶装置ビットマップ格納領域31に格納された記憶装置ビットマップ22aの少なくとも一部である。ビットマップ領域22が保持する記憶装置ビットマップ22a及びメモリビットマップ22bの詳細については、後述する。

【0043】

SSD(格納部)30は、種々のデータやプログラムを格納する記憶装置であり、記憶装置ビットマップ格納領域31を含む。記憶装置ビットマップ格納領域31は、上述したHDD5の記憶装置ビットマップ格納領域53に格納されるもの同一の記憶装置ビットマップ22aを格納する。なお、図1には、SSD30及びHDD5の双方が記憶装置ビットマップ22aを格納する例を示しているが、SSD30及びHDD5のうちのいずれか一方のみが記憶装置ビットマップ22aを格納することとしてもよい。また、SSD30又は/及びHDD5の記憶装置ビットマップ格納領域53に代えて、CPU10からアクセス可能な他の記憶装置が用いられてもよい。さらに、コントローラモジュール4には、SSD30に代えて、HDD等がそなえられてもよい。

【0044】

〔1-3〕処理部の説明

次に、CPU10が処理部として実現する各機能について説明する。

割当処理部14は、HDD5のデータ領域52bを複数の第1サイズのデータブロックの記憶領域に割り当て、複数のデータ領域52bに同じ第1サイズのデータブロックを割り当てる処理を行なう。つまり、割当処理部14は、上述のように、例えば数十セクタ～数百セクタの単位で、複数のA領域52b-1、複数のB領域52b-2、複数のA領域52b-1、...という形で、データ領域52bを交互に割り当てる(図1参照)。

【0045】

なお、割当処理部14によるLBAとHDD5の一つのセクタとの間でのアドレスの変換は、既知の種々の手法を用いて実現可能である。本実施形態においては、このような手法に対して、一つのLBAについて対応する複数(二つ)のセクタを得るように論理を変更することで、LBAとHDD5の複数(二つ)のセクタとの間でのアドレスの変換を実現することができる。

【0046】

記憶装置ビットマップ管理部(管理部)13は、HDD5にホスト装置3からのデータブロックをA領域52b-1及びB領域52b-2のいずれか一つに排他的に保存させるように、複数のデータ領域52bにおけるデータ保存状態を管理する。記憶装置ビットマップ管理部13によるデータ保存状態の管理は、メモリ20のビットマップ領域22に保持された記憶装置ビットマップ22a及びメモリビットマップ22bに基づいて行なわれる。例えば、記憶装置ビットマップ管理部13は、ビットマップ更新部13a、ビットマップ読出部13b、及び、ビットマップ書込部13cを備える。

【0047】

ビットマップ更新部13aは、メモリ20に保持されたデータブロックが、後述する書込制御部11により一のデータ領域52bへライトバックにより書き込まれると、メモリビットマップ22bに基づき記憶装置ビットマップ22aを作成又は更新する。

具体的には、ビットマップ更新部13aは、ホスト装置3からユーザデータ領域21に書き込まれたデータブロックが、一のデータ領域52bへ書き込まれると、該データブロックに対して割り当てられた複数の記憶領域のデータ保存状態を更新する。

【0048】

より具体的に、ビットマップ更新部13aは、書き込み対象のデータブロックに対して

10

20

30

40

50

割り当てられた複数の記憶領域のデータ保存状態について、以下の二つの条件を満たすように記憶装置ビットマップ 2 2 a を書き換える。

- ・上記複数の記憶領域のうち、一のデータ領域 5 2 b 内で書き込み対象のデータブロックに対して割り当てられた記憶領域はデータブロックのデータが書き込まれると有効を示す。

- ・上記複数の記憶領域のうち、他の全てのデータ領域 5 2 b 内で書き込み対象のデータブロックに対して割り当てられた記憶領域は無効を示す。

【 0 0 4 9 】

例えば、ビットマップ更新部 1 3 a は、ホストブロックが書き込まれたデータ領域 5 2 b の記憶装置ビットマップ 2 2 a に対して、メモリビットマップ 2 2 b において “ 1 ” が設定されたデータブロックに対応する位置のビットを “ 1 ” に更新する。

なお、ビットマップ更新部 1 3 a は、HDD 5 がフォーマットされた場合等、対象となる記憶装置ビットマップ 2 2 a が存在しない場合には、対象となる記憶装置ビットマップ 2 2 a を新たに作成してビットマップ領域 2 2 に格納する。

【 0 0 5 0 】

なお、HDD 5 がフォーマットされる場合、第 1 グループ 5 2 a - 1 及び第 2 グループ 5 2 a - 2 の全データ領域 5 2 b に、有効なデータ（例えば “ 0 ”）が設定される。この場合、全データ領域 5 2 b 内に有効なデータが存在するが、ビットマップ更新部 1 3 a は、第 1 グループ 5 2 a - 1 及び第 2 グループ 5 2 a - 2 のいずれか一方のグループについて、全てのビットが “ 1 ”（有効）の記憶装置ビットマップ 2 2 a を作成する。

【 0 0 5 1 】

ビットマップ読出部 1 3 b は、SSD 3 0 又は HDD 5 等から記憶装置ビットマップ 2 2 a のうちの、CPU 1 0 から指示されたデータ領域 5 2 b に係る部分を読み出し、メモリ 2 0 のビットマップ領域 2 2 に展開する。

なお、ビットマップ読出部 1 3 b は、SSD 3 0 等から読み出す記憶装置ビットマップ 2 2 a の部分を、CPU 1 0 からの指示ではなく、CPU 1 0 内部又はチャネルアダプタ 6 を監視することで推定してもよい。つまり、ビットマップ読出部 1 3 b は、監視により得られたホスト装置 3 からの書込 / 読出コマンドから、CPU 1 0 による処理対象となるデータ領域 5 2 b を推定してもよい。処理対象となるデータ領域 5 2 b を推定することにより、ビットマップ読出部 1 3 b は、CPU 1 0 が要求する記憶装置ビットマップ 2 2 a を予めビットマップ領域 2 2 に格納しておくことができ、コントローラモジュール 4 の処理性能を向上させることができる。

【 0 0 5 2 】

ビットマップ書込部 1 3 c は、ビットマップ領域 2 2 に保持された記憶装置ビットマップ 2 2 a を、所定のタイミングで SSD 3 0 又は HDD 5 等へ書き込み、記憶装置ビットマップ格納領域 3 1 又は 5 3 内の記憶装置ビットマップ 2 2 a を更新する。なお、所定のタイミングの例としては、以下の (i) ~ (iii) のうちの一つ又はこれらの組み合わせが挙げられる。

(i) 定期的なタイミング

(ii) 書込制御部 1 1 により HDD 5 へライトバックが行なわれたタイミング

(iii) ビットマップ領域 2 2 内の記憶装置ビットマップ 2 2 a のサイズが所定サイズ以上になったタイミング

このように、記憶装置ビットマップ管理部 1 3 は、複数のデータ領域 5 2 b のうちのいずれか一のデータ領域 5 2 b 内でデータブロックに対して割り当てられた記憶領域が、他の全てのデータ領域内で該データブロックに対して割り当てられた記憶領域との間で排他的に有効を示すように、データ保存状態を管理する。なお、記憶装置ビットマップ管理部 1 3 は、各データ領域 5 2 b に割り当てられた複数のデータブロックの記憶領域について、データブロックごとに、上記の管理を行なう。

【 0 0 5 3 】

メモリビットマップ管理部 1 5 は、ホスト装置 3 からのユーザデータ（書き込み対象の

ユーザデータ)が、HDD5内の対応するデータ領域52bのうちのいずれのデータブロックに書き込まれるか否かを示すメモリビットマップ22bを管理する。

メモリビットマップ(第2管理情報)22bは、書き込み対象のデータブロックに対して割り当てられた複数のデータ領域52bにおける書き込み対象のデータブロックが保存される領域を示す情報である。図2に示すLBA#00~#07に着目して説明すると、例えば、LBA#03及び#04のデータがHDD5のセクタ#0-0又は#0-1に書き込むべきユーザデータとしてホスト装置3から入力された場合、メモリビットマップ22bは以下ようになる。

【0054】

LBA#00~#07:00011000

10

なお、上記のメモリビットマップ22bは、1ビットが512バイトの1ホストブロックに相当し、左からLBAの昇順で配置される。また、メモリビットマップ22bにおいて、“1”は有効つまり対応する位置のホストブロックのデータが書き込み対象であることを示し、“0”は無効つまり対応する位置のホストブロックのデータが書き込み対象ではないことを示す。

【0055】

メモリビットマップ管理部15は、ホスト装置3からの書込コマンドに応じて、上述したメモリビットマップ22bを作成し、メモリ20のビットマップ領域22に格納する。

ここで、書き込み対象のユーザデータは、ユーザデータ領域21の任意の位置に格納される。従って、メモリビットマップ管理部15は、ユーザデータ領域21における書き込み対象のユーザデータが格納された位置と、メモリビットマップ22bとを対応付けて管理することが好ましい。

20

【0056】

なお、メモリビットマップ管理部15は、ユーザデータ領域21に保持されたユーザデータがHDD5へライトバックされると、当該ユーザデータに対応するメモリビットマップ22bを破棄、又は、ALL“0”に更新する。

〔1-3-1〕書込制御部の説明

書込制御部11は、ホスト装置3からの書込コマンドに応じて書込データをメモリ20に一旦留め、ライトバック等により非同期でHDD5へ書き込む制御を行なう。例えば、書込制御部11は、書込領域判定部11a及び書込処理部11bを備える。

30

【0057】

書込領域判定部11aは、ビットマップ領域22に保持された、書き込み対象のデータブロックに対して割り当てられた複数のデータ領域52bの記憶装置ビットマップ22aに基づき、複数のデータ領域52bの中から一のデータ領域52bを判定する。ここで、一のデータ領域52bとは、HDD5においてデータブロックの書き込みに伴うデータブロックの読み出し、つまり、RMWが可能な限り発生しないデータ領域52bである。

【0058】

なお、書込領域判定部11aは、RMWの発生しないデータ領域52bを判定できない場合がある。この場合、HDD5においてRMWが発生することになるが、以下の処理により、今回又は次回以降のRMWによるHDD5の性能低下への影響を抑えたり、RMWの発生しないデータ領域52bを判定し易くしたりすることができる。

40

例えば、書込領域判定部11aは、RMWの発生しないデータ領域52bを判定できない場合、所定の条件に応じて、複数のデータ領域52bからいずれか一のデータ領域52bを選択する。所定の条件としては、HDD5の処理負荷が所定の閾値以上であることや、ホスト装置3から書込/読出コマンドが所定数以上発行されていること等、HDD5の性能低下を引き起こす可能性が高まったことを条件とすることができる。

【0059】

書込領域判定部11aは、HDD5の性能低下を引き起こす可能性が高まった場合、RMWにおいてHDD5によるシーク処理及びマージ処理(図14中(3)参照)の処理量が少なくなるデータ領域52bを選択する。一方、書込領域判定部11aは、HDD5の

50

性能低下を引き起こす可能性が低い場合、今回又は次回以降 R M W の発生しないデータ領域 5 2 b を判定し易くするために、記憶装置ビットマップ 2 2 a が複数のデータ領域 5 2 b 間で偏るようなデータ領域 5 2 b を選択する。

【 0 0 6 0 】

書込領域判定部 1 1 a による一のデータ領域 5 2 b の判定の手法については、後述する。

書込処理部 1 1 b は、書き込み対象のデータブロックのデータを含む第 2 サイズの書込用データブロック（書込用データ）を生成し、生成した書込用データブロックを、書込領域判定部 1 1 a により判定された一のデータ領域 5 2 b へ書き込む処理を行なう。

【 0 0 6 1 】

具体的には、書込処理部 1 1 b は、ユーザデータ領域 2 1 において、書き込み対象のユーザデータ（データブロック）にダミーデータを付加して 4 0 9 6 バイト単位に丸めることで、書込用データブロックを生成する。より具体的に、書込処理部 1 1 b は、メモリビットマップ 2 2 b とユーザデータ領域 2 1 内のユーザデータの格納位置とに基づき、書き込み対象のユーザデータを含む 4 0 9 6 バイト分のユーザデータ領域 2 1 内のデータ（書込用データブロック）を取得する。取得されたデータには、書き込み対象のユーザデータ（データブロック）と、ユーザデータ領域 2 1 のアドレス上における、書き込み対象のユーザデータの前後少なくとも一方のダミーデータ（不要データ）とが含まれる。

【 0 0 6 2 】

なお、ダミーデータは、既に書込制御部 1 1 により H D D 5 へライトバックが行なわれたデータブロックである。従って、書込処理部 1 1 b は、書込用データブロックの生成にあたり、ユーザデータ領域 2 1 に残存する不要データブロックを流用することができるため、ダミーデータを生成してユーザデータに付加するよりも高速に書込用データブロックを生成することができる。なお、コントローラモジュール 4 は、メモリビットマップ 2 2 b によりユーザデータとダミーデータとを区別することができる。

【 0 0 6 3 】

図 2 に示す L B A # 0 0 ~ # 0 7 に着目し、例えば、ホスト装置 3 から書き込み対象のユーザデータとして L B A # 0 3 及び # 0 4 のデータがメモリ 2 0 に書き込まれた場合について説明する。書込処理部 1 1 b は、ユーザデータ領域 2 1 における L B A # 0 3 及び # 0 4 の書き込み対象のデータブロックと、L B A # 0 0 ~ # 0 2 及び # 0 5 ~ # 0 7 に

対応するユーザデータ領域 2 1 におけるデータブロックと、をまとめて書込用データブロックとして取得する。つまり、書込処理部 1 1 b は、ユーザデータ領域 2 1 に保持された書き込み対象のデータブロックを含む二以上のデータブロックを用いて書込用データブロックを生成するのである。

【 0 0 6 4 】

このように、書込処理部 1 1 b によれば、H D D 5 において、4 0 9 6 バイト単位でデータブロックが書き込まれるため R M W が発生せず、従来の 5 1 2 バイト / セクタの H D D と同等な書込性能を維持することができる。また、書込用データブロックの生成にあたりユーザデータ領域 2 1 内の実際のデータを用いるため、4 0 9 6 バイトの書込用データブロックを容易に生成することができる。

【 0 0 6 5 】

なお、書込処理部 1 1 b は、ホスト装置 3 からのデータブロックに “ 0 ” 又は “ 1 ” の任意のダミーデータを付加することで書込用データブロックを生成してもよい。

また、上述したように、書込領域判定部 1 1 a により R M W の発生しないデータ領域 5 2 b が判定されない場合がある。この場合、書込処理部 1 1 b は、書き込み対象のデータブロックに対して割り当てられた複数のデータ領域 5 2 b の記憶装置ビットマップ 2 2 a に基づいて、書き込み対象のデータブロックを含む所定のサイズの書込用データブロックを生成する。そして、書込処理部 1 1 b は、生成した所定のサイズの書込用データブロックを、書込領域判定部 1 1 a により選択された一のデータ領域 5 2 b へ書き込む。

【 0 0 6 6 】

なお、書込領域判定部 11a は、ホスト装置 3 からのユーザデータが第 2 サイズである 4096 バイト単位である場合には、RMW が発生しないため、データ領域 52b の判定処理を行わずに対応するデータ領域 52b のうちの任意の一方を選択する。そして、書込処理部 11b は、第 2 サイズであるユーザデータを、書込領域判定部 11a が選択した一のデータ領域 52b へ書き込む。

【0067】

換言すれば、書込領域判定部 11a は、ホスト装置 3 からのユーザデータが、RMW が発生する条件に該当する場合に、記憶装置ビットマップ 22a に基づく一のデータ領域 52b の判定を行なうのである。なお、RMW が発生する条件として、ホスト装置 3 からのユーザデータが 4096 バイトに満たない（又は 4096 バイトの整数倍のサイズではない）場合が挙げられる。また、RMW が発生する他の条件として、4096 バイトの整数倍のサイズであるユーザデータの開始アドレスがデータ領域 52b の開始アドレスからずれている場合等も挙げられる。

【0068】

なお、ビットマップ更新部 13a は、書込処理部 11b により書込用データブロックがデータ領域 52b に書き込まれると、上述のように、書込用データブロックに含まれる書き込み対象のデータブロックについて、記憶装置ビットマップ 22a を更新する。

次に、図 3～図 9 を参照して、書込制御部 11 の動作例を説明する。図 3～図 8 は、それぞれ、図 1 に示す処理部による動作例を説明するための図であり、図 9 (a) 及び図 9 (b) は、それぞれ、図 7 及び図 8 に示す動作例における HDD 5 でのマージ量を説明するための図である。なお、図 3～図 9 において、枠で囲われた“記憶領域”及び“ユーザデータ領域”内の数字はデータブロックの世代数を示し、英字は当該データブロックが意味のない（無効な）データを設定されたブロックであることを示す。また、図 3～図 9 において、記憶装置ビットマップ 22a は、A 領域 52b - 1（第 1 グループ 52a - 1）についてのビットマップ情報であるものとする。

【0069】

〔1 - 3 - 1 - 1〕A 領域又は B 領域の記憶装置ビットマップが ALL “1” である場合

書込領域判定部 11a は、データブロックを書き込むデータ領域 52b の判定において、対応する A 領域 52b - 1 及び B 領域 52b - 2 のいずれか一方の記憶装置ビットマップ 22a のデータ保存状態が ALL “1” であるか否かを判定する。A 領域 52b - 1 及び B 領域 52b - 2 のいずれか一方が ALL “1” と判定すると、書込領域判定部 11a は、ALL “1” ではないデータ領域 52b を、データブロックを書き込むべき一のデータ領域 52b と判定する。

【0070】

換言すれば、書込領域判定部 11a は、一のデータ領域 52b として、書き込み対象のデータブロックに対して割り当てられた複数のデータ領域 52b の記憶装置ビットマップ 22a のデータ保存状態が ALL “0” のデータ領域 52b を選択するのである。

一例として、図 3 を参照して説明する。前提として、HDD 5 の記憶領域 52 はフォーマットされた直後であり、全てのデータ領域 52b には ALL “0” のデータが記憶されているものとする。図 3 の (1) に示すように、フォーマット直後の記憶装置ビットマップ 22a には、ビットマップ更新部 13a により A 領域 52b - 1 について ALL “1” が設定されている。この場合、図 3 の (2) に示すように、B 領域 52b - 2 についての記憶装置ビットマップ 22a は、ALL “0” である。

【0071】

図 3 の (3) において、ホスト装置 3 により、LBA #00～#03 について世代 “1” のデータブロックの書込コマンドが発行されると、メモリビットマップ管理部 15 により、先頭 4 ビットが “1” のメモリビットマップ 22b が生成される。

書込領域判定部 11a は、書込コマンドに応じて、書き込み対象の LBA に対応するセクタ #0 - 0（A 領域 52b - 1）の記憶装置ビットマップ 22a を参照し、全てのビッ

10

20

30

40

50

トが“1”であるか否かを判定する。図3の(2)に示す例では、A領域52b-1の記憶装置ビットマップ22aはALL“1”であるため、書込領域判定部11aは、A領域52b-1とは異なるB領域52b-2を、データブロックを書き込むべき一のデータ領域52bと判定する。

【0072】

また、書込処理部11bは、書き込み対象のLBA#00~#03のデータブロック(2048バイトのショートブロック)を、書込領域判定部11aが判定したB領域52b-2に書き込む。このとき、LBA#00~#03の4ブロックのみをB領域52b-2にライトバックすると、HDD5においてRMWが発生する。そこで、書込処理部11bは、メモリビットマップ22bに基づき、ユーザデータ領域21から、書き込み対象のLBA#00~#03のデータブロックと、ダミーデータであるLBA#04~#07のデータブロックとを生成(取得)する。そして、書込処理部11bは、生成した4096バイトの書込用データブロックを、書込領域判定部11aが判定したデータ領域52bに書き込む(図3の(4)参照)。

【0073】

さらに、書込処理部11bが書込用データブロックの書き込みを行なうと、ビットマップ更新部13aがメモリビットマップ22bに基づき記憶装置ビットマップ22aを更新する。

なお、書込領域判定部11aは、A領域52b-1の記憶装置ビットマップ22aを参照し、全てのビットが“1”であるか否かを判定するものとして説明したが、これに限定されるものではない。例えば、書込領域判定部11aは、A領域52b-1の記憶装置ビットマップ22aの全てのビットが“0”であるか否かを判定してもよい。この場合、“0”である場合には、書込領域判定部11aは、A領域52b-1を、データブロックを書き込むべき一のデータ領域52bと判定する。

【0074】

〔1-3-1-2〕メモリビットマップと、A領域又はB領域の記憶装置ビットマップとの論理和がALL“1”である場合

書込領域判定部11aは、A領域52b-1及びB領域52b-2のいずれか一方に書き込み対象のデータブロックが保存されると、一方のデータ領域52bにおける全ての記憶装置ビットマップ22aのデータ保存状態がALL“1”になるか否かを判定する。データ保存状態がALL“1”になると判定すると、書込領域判定部11aは、ALL“1”になるデータ領域52bとは異なる他のデータ領域52bを、データブロックを書き込むべき一のデータ領域52bと判定する。

【0075】

換言すれば、書込領域判定部11aは、データ領域52bの記憶装置ビットマップ22aにおいて“1”である全てのビット位置が、書き込み対象のデータブロックのメモリビットマップ22bにおいて“1”である全てのビット位置に含まれるか否かを判定する。つまり、書込領域判定部11aは、前記書き込み対象のデータブロックに対して割り当てられたデータ領域52bにおけるデータ保存状態が有効な全てのデータブロックの領域が、データ領域52bにおいて書き込み対象のデータブロックが保存される全ての領域に含まれるか否かを判断する。そして、書込領域判定部11aは、含まれると判定した場合に、当該データ領域52bを、データブロックを書き込むべき一のデータ領域52bと判定する。

【0076】

一例として、図4を参照して説明する。前提として、HDD5の記憶領域52はランダムライトが行なわれた後であり、複数のデータ領域52bには、飛び飛びでデータが書き込まれた状態であるものとする。図4の(1)に示すように、記憶装置ビットマップ22aには、ビットマップ更新部13aによりA領域52b-1について飛び飛びで“1”が設定されている。

【0077】

図4の(2)において、ホスト装置3により、LBA#00~#07のデータブロックの書込コマンドが発行されると、メモリビットマップ管理部15により、先頭8ビットが“1”のメモリビットマップ22bが生成される。

書込領域判定部11aは、書込コマンドに応じて、書き込み対象のLBAに対応するセクタ#0-0(A領域52b-1)の記憶装置ビットマップ22aを参照し、メモリビットマップ22bとの論理和を取った場合に、全てのビットが“1”になるか否かを判定する。図4の(2)に示す例では、メモリビットマップ22bだけで8ホストブロック全てが“1”であり、バウンダリに合った4096バイトのユーザデータであるため、書込領域判定部11aは、A領域52b-1及びB領域52b-2を選択するまでもない。例えば、書込領域判定部11aがA領域52b-1を選択した場合、書込処理部11bにより、4096バイトである書き込み対象のデータブロックがA領域52b-1に書き込まれる(図4の(3)参照)。

10

【0078】

また、他の例を、図5を参照して説明する。前提として、HDD5の記憶領域52はランダムライトが行なわれた後であり、複数のデータ領域52bには、1024バイトで離散したデータが書き込まれた状態であるものとする。図5の(1)に示すように、記憶装置ビットマップ22aには、ビットマップ更新部13aによりA領域52b-1について2ブロックごとに飛び飛びで“1”が設定されている。

【0079】

図5の(2)において、ホスト装置3により、LBA#00、#01、#04、#05のデータブロックの書込コマンドが発行されると、メモリビットマップ管理部15により、2ビットごとに飛び飛びで“1”であるメモリビットマップ22bが生成される。

20

書込領域判定部11aは、書込コマンドに応じて、書き込み対象のLBAに対応するセクタ#0-0(A領域52b-1)の記憶装置ビットマップ22aを参照し、メモリビットマップ22bとの論理和を取った場合に、全てのビットが“1”になるか否かを判定する。図5の(2)に示す例では、A領域52b-1の記憶装置ビットマップ22aとメモリビットマップ22bとの論理和がALL“1”になるため、書込領域判定部11aは、A領域52b-1とは異なるB領域52b-2を、データブロックを書き込むべきデータ領域52bと判定する。

【0080】

30

また、書込処理部11bは、メモリビットマップ22bに基づき、書き込み対象のLBA#00、#01、#04、#05のデータブロックと、ダミーデータであるLBA#02、#03、#06、#07のデータブロックとを生成(取得)する。そして、書込処理部11bは、生成した4096バイトの書込用データブロックを、書込領域判定部11aが判定したデータ領域52bに書き込む(図5の(3)参照)。

【0081】

さらに、他の例を、図6を参照して説明する。前提として、ホスト装置3がメインフレームであり、ホスト装置3から指定される書き込み対象のアドレスがLBAではなくCount Key Data(CKD)であるものとする。なお、CKDは、書き込み対象のデータブロックのサイズがレコードという固定サイズであり(図6の(1)“R1”~“R4”参照)、1レコードは、512バイトよりも大きく、且つ、512バイトの整数倍ではないものとする。また、コントローラモジュール4は、ホスト装置3から指定されるCKDをLBAに変換して、書き込み対象のレコード(書き込み対象のデータブロック)をLBAに対応付ける。さらに、コントローラモジュール4は、レコードを512バイト単位で区切ったときの512バイト未満の部分にダミーデータを付加して、当該部分のバウンダリを512バイトに合わせる。なお、図6に示す例では、レコードのサイズは、ダミーデータが付加されて5データブロック(2560バイト)に調整される。

40

【0082】

さらに、前提として、HDD5の記憶領域52はフォーマットされた直後であり、全てのデータ領域52bにはALL“0”のデータが記憶されているものとする。図6の(2

50

）に示すように、フォーマット直後の記憶装置ビットマップ 2 2 a には、ビットマップ更新部 1 3 a により A 領域 5 2 b - 1（第 1 グループ 5 2 a - 1）について A L L “ 1 ” が設定されている。

【 0 0 8 3 】

上述した前提において、ホスト装置 3 により、R 2 について世代 “ 1 ” のレコードの書込コマンドが発行されると、メモリ 2 0 のユーザデータ領域 2 1 に、世代 “ 1 ” の R 2 のデータが格納される（図 6 の（ 3 ）参照）。また、メモリビットマップ管理部 1 5 により、R 2 に対応するビット（ 6 ビット目～ 1 0 ビット目のビット）が “ 1 ” のメモリビットマップ 2 2 b が生成される。

【 0 0 8 4 】

書込領域判定部 1 1 a は、書込コマンドに応じて、書き込み対象のレコードに対応するセクタ # 0 - 0 及び # 1 - 0（A 領域 5 2 b - 1）の記憶装置ビットマップ 2 2 a を参照し、それぞれのセクタにおいて全てのビットが “ 1 ” であるか否かを判定する。図 6 の（ 4 ）に示す例では、セクタ # 0 - 0 及び # 1 - 0 のいずれの記憶装置ビットマップ 2 2 a も A L L “ 1 ” である。従って、書込領域判定部 1 1 a は、A 領域 5 2 b - 1 とは異なる B 領域 5 2 b - 2 であるセクタ # 0 - 1 及び # 1 - 1 を、それぞれレコードを書き込むべき一のデータ領域 5 2 b と判定する。

【 0 0 8 5 】

また、書込処理部 1 1 b は、書き込み対象の R 2 のレコードを、書込領域判定部 1 1 a が判定した B 領域 5 2 b - 2 であるセクタ # 0 - 1 及び # 1 - 1 に書き込む。このとき、書込処理部 1 1 b は、メモリビットマップ 2 2 b に基づき、ユーザデータ領域 2 1 から、書き込み対象の R 2 のレコードと、ダミーデータであるデータブロック（図 6 の（ 3 ）セクタ # 0 - 0 ～ # 1 - 0 に対応するデータ “ x ” 参照）とを生成（取得）する。そして、書込処理部 1 1 b は、生成した 4 0 9 6 バイト × 2 の書込用データブロックを、書込領域判定部 1 1 a が判定したデータ領域 5 2 b に書き込む（図 6 の（ 4 ）参照）。

【 0 0 8 6 】

次いで、ホスト装置 3 により、R 1 について世代 “ 1 ” のレコードの書込コマンドが発行されると、メモリ 2 0 のユーザデータ領域 2 1 に、世代 “ 1 ” の R 1 のデータが格納される（図 6 の（ 5 ）参照）。また、メモリビットマップ管理部 1 5 により、R 1 に対応するビット（ 1 ビット目～ 5 ビット目のビット）が “ 1 ” のメモリビットマップ 2 2 b が生成される。

【 0 0 8 7 】

書込領域判定部 1 1 a は、書込コマンドに応じて、書き込み対象のレコードに対応するセクタ # 0 - 0（A 領域 5 2 b - 1）の記憶装置ビットマップ 2 2 a を参照し、それぞれのセクタにおいて全てのビットが “ 1 ” であるか否かを判定する。図 6 の（ 4 ）に示す例では、セクタ # 0 - 0 の記憶装置ビットマップ 2 2 a は A L L “ 1 ” ではない。

そこで、書込領域判定部 1 1 a は、書き込み対象のレコードに対応するセクタ # 0 - 0 及び # 0 - 1（A 領域 5 2 b - 1）の記憶装置ビットマップ 2 2 a を参照し、メモリビットマップ 2 2 b との論理和を取った場合に、全てのビットが “ 1 ” になるか否かを判定する。図 6 の（ 5 ）に示す例では、B 領域 5 2 b - 2 の記憶装置ビットマップ 2 2 a とメモリビットマップ 2 2 b との論理和が A L L “ 1 ” になるため、書込領域判定部 1 1 a は、B 領域 5 2 b - 1 とは異なる A 領域 5 2 b - 2 を、データブロックを書き込むべき一のデータ領域 5 2 b と判定する。

【 0 0 8 8 】

また、書込処理部 1 1 b は、書き込み対象の R 1 のレコードを、書込領域判定部 1 1 a が判定した A 領域 5 2 b - 1 であるセクタ # 0 - 0 に書き込む。このとき、書込処理部 1 1 b は、メモリビットマップ 2 2 b に基づき、ユーザデータ領域 2 1 から、書き込み対象の R 1 のレコードと、ダミーデータであるデータブロック（図 6 の（ 5 ）セクタ # 0 - 0 に対応するデータ “ x ” 参照）とを生成（取得）する。そして、書込処理部 1 1 b は、生成した 4 0 9 6 バイトの書込用データブロックを、書込領域判定部 1 1 a が判定したデー

10

20

30

40

50

タ領域 5 2 b に書き込む（図 6 の（ 6 ）参照）。

【 0 0 8 9 】

図 6 を用いて説明したように、本実施形態に係るストレージ装置 2 は、ホスト装置 3 がメインフレームである場合でも、R M W の発生を抑止し、H D D 5 の性能低下を軽減させることができる。特に、レコードサイズが第 2 サイズである 4 0 9 6 バイト（ 1 セクタ）よりも大きい場合には、一つのセクタに属するレコードは最大でも二つになる。これは、書込領域判定部 1 1 a が、A 領域 5 2 b - 1 又は B 領域 5 2 b - 2 から、データブロックを書き込むべき一のデータ領域 5 2 b を確実に判定できることを意味する。すなわち、レコードサイズが第 2 サイズよりも大きい場合には、R M W が発生しなくなるため、R M W に起因する H D D 5 の性能低下を無くすることができる。

10

【 0 0 9 0 】

なお、図 4 ～ 図 6 の例においても、書込処理部 1 1 b が書込用データブロックの書き込みを行なうと、ビットマップ更新部 1 3 a がメモリビットマップ 2 2 b に基づき記憶装置ビットマップ 2 2 a を更新する。

また、書込領域判定部 1 1 a は、A 領域 5 2 b - 1 又は B 領域 5 2 b - 2 の記憶装置ビットマップ 2 2 a とメモリビットマップ 2 2 b との論理和が A L L “ 1 ” になるか否かを判定するものとして説明したが、これに限定されるものではない。

【 0 0 9 1 】

例えば、書込領域判定部 1 1 a は、上述のように、データ領域 5 2 b の記憶装置ビットマップ 2 2 a において “ 1 ” である全てのビット位置が、メモリビットマップ 2 2 b において “ 1 ” である全てのビット位置に含まれるか否かを判定してもよい。この場合、書込領域判定部 1 1 a は、含まれると判定したデータ領域 5 2 b を、データブロックを書き込むべき一のデータ領域 5 2 b と判定する。

20

【 0 0 9 2 】

〔 1 - 3 - 1 - 3 〕その他の場合

上述した例では、書込領域判定部 1 1 a が、R M W の発生しないデータ領域 5 2 b を判定できる場合について説明したが、ここでは、R M W の発生しないデータ領域 5 2 b を判定できない場合について説明する。

一例として、図 7 及び図 8 を用いて説明する。前提として、H D D 5 の記憶領域 5 2 は書込処理が行なわれた後であり、複数のデータ領域 5 2 b には、図 7 の（ 1 ）及び図 8 の（ 1 ）に示すようなデータが書き込まれた状態であるものとする。図 7 の（ 1 ）及び図 8 の（ 1 ）に示すように、記憶装置ビットマップ 2 2 a には、ビットマップ更新部 1 3 a により A 領域 5 2 b - 1 について飛び飛びで “ 1 ” が設定されている。

30

【 0 0 9 3 】

図 7 の（ 2 ）及び図 8 の（ 2 ）において、ホスト装置 3 により、L B A # 0 7 のデータブロックの書込コマンドが発行されると、メモリビットマップ管理部 1 5 により、最終 1 ビットが “ 1 ” のメモリビットマップ 2 2 b が生成される。

書込領域判定部 1 1 a は、書込コマンドに応じて、R M W の発生しないデータ領域 5 2 b の判定を試みる。しかし、書き込み対象の L B A に対応するセクタ # 0 - 0 （ A 領域 5 2 b - 1 ）は、全ビットが “ 1 ” ではなく、セクタ # 0 - 0 の記憶装置ビットマップ 2 2 a とメモリビットマップ 2 2 b との論理和も全ビットが “ 1 ” にはならない。

40

【 0 0 9 4 】

このような場合には R M W が発生するが、書込領域判定部 1 1 a は、既述のように所定の条件に応じて、A 領域 5 2 b - 1 及び B 領域 5 2 b - 2 のいずれか一方のデータ領域 5 2 b を選択する。

図 7 に示す例では、バックエンド負荷が低く、H D D 5 の性能低下を引き起こす可能性が低い場合について説明する。

【 0 0 9 5 】

この場合、書込領域判定部 1 1 a は、H D D 5 において R M W によりマージされる量が多いデータ領域 5 2 b を選択する。マージ量が多いということは、データ領域 5 2 b 中に

50

有効データが多くなることを意味する。つまり、書込領域判定部 1 1 a は、A 領域 5 2 b - 1 又は B 領域 5 2 b - 2 のいずれかに有効データが多いデータ領域 5 2 b を用意しておくのである。これにより、次回以降の書き込みにおいて、ホスト装置 3 から書き込まれるホストブロック数が少なくても、ライトバックの際に、A 領域 5 2 b - 1 の記憶装置ビットマップ 2 2 a とメモリビットマップ 2 2 b との論理和が全ビット “ 1 ” になる可能性が高くなる。

【 0 0 9 6 】

例えば、図 7 の (3) に示すように、書込領域判定部 1 1 a は、H D D 5 において R M W による読出量が多いが 1 セクタ内に有効データが多くなるように、データブロックを書き込むべき一のデータ領域 5 2 b として B 領域 5 2 b - 2 を選択する。

10

また、書込処理部 1 1 b は、書き込み対象の L B A # 0 7 のデータブロックを、書込領域判定部 1 1 a が判定した B 領域 5 2 b - 2 に書き込む (図 7 の (4) 参照) 。

【 0 0 9 7 】

さらに、H D D 5 は、図 9 (a) に示すように、R M W により記憶領域 5 2 のセクタ # 0 - 1 (B 領域 5 2 b - 2) から、7 データブロック分のデータを読み出し、メモリ 5 1 で書込処理部 1 1 b からの 1 データブロックとのマージ処理を行なう。そして、H D D 5 は、マージ結果をセクタ # 0 - 1 へ書き込み、処理を終了する。

なお、書込処理部 1 1 b が書込用データブロックの書き込みを行なうと、ビットマップ更新部 1 3 a が記憶装置ビットマップ 2 2 a を更新する。

【 0 0 9 8 】

20

上記の処理により、セクタ # 0 - 1 では、L B A # 0 1 ~ # 0 7 に対応する 7 データブロックが有効なデータとなる。

一方、図 8 に示す例では、バックエンド負荷が高く、H D D 5 の性能低下を引き起こす可能性が高い場合について説明する。

この場合、書込領域判定部 1 1 a は、バックエンド負荷をこれ以上上げないため、マージ量が少ない面を選択する。

【 0 0 9 9 】

例えば、図 8 の (3) に示すように、書込領域判定部 1 1 a は、H D D 5 において R M W による読出量が少ない A 領域 5 2 b - 1 を、データブロックを書き込むべき一のデータ領域 5 2 b に選択する。

30

また、書込処理部 1 1 b は、書き込み対象の L B A # 0 7 のデータブロックを、書込領域判定部 1 1 a が判定した B 領域 5 2 b - 2 に書き込む。このとき、書込処理部 1 1 b は、メモリビットマップ 2 2 b に基づき、ユーザデータ領域 2 1 から、書き込み対象の L B A # 0 7 のデータブロックと、ダミーデータである L B A # 0 1 ~ # 0 6 のデータブロックとを生成 (取得) する。そして、書込処理部 1 1 b は、生成した 7 データブロック分の書込用データブロックを、書込領域判定部 1 1 a が判定したデータ領域 5 2 b に書き込む (図 8 の (4) 参照) 。

【 0 1 0 0 】

さらに、H D D 5 は、図 9 (b) に示すように、R M W により記憶領域 5 2 のセクタ # 0 - 0 (A 領域 5 2 b - 1) から、1 データブロック分のデータを読み出し、メモリ 5 1 で書込処理部 1 1 b からの 7 データブロックとのマージ処理を行なう。そして、H D D 5 は、マージ結果をセクタ # 0 - 0 へ書き込み、処理を終了する。

40

なお、書込処理部 1 1 b が書込用データブロックの書き込みを行なうと、ビットマップ更新部 1 3 a が記憶装置ビットマップ 2 2 a を更新する。

【 0 1 0 1 】

上記の処理により、図 9 (b) に示すように、H D D 5 において R M W により記憶領域 5 2 のセクタ # 0 - 0 から読み出すデータは、1 データブロック分で済む。従って、図 9 (a) の場合と比較して、H D D 5 の性能劣化を低減させることができる。

以上のように、本実施形態に係るコントローラモジュール 4 は、ホスト装置 3 から見たデータブロックに対して、H D D 5 内に複数倍 (例えば二倍) のデータブロックを持つよ

50

うに制御を行なう。そして、書込制御部 11 は、書き込み対象のデータブロックを HDD 5 へ書き込む際に、可能な限り RMW が発生しないデータ領域 52b を選択し、書き込みを行なうことで、性能の低下を抑えるように制御を行なう。

【0102】

〔1-3-2〕読出制御部の説明

図 1 の説明に戻り、読出制御部 12 は、ホスト装置 3 からの読出要求に応じて、HDD 5 から読出データを読み出し、ホスト装置 3 へ渡す制御を行なう。例えば、読出制御部 12 は、読出領域判定部 12a、読出処理部 12b、及び、データマージ処理部 12c を備える。

【0103】

読出領域判定部 12a は、読み出し対象のデータを形成する第 1 サイズの一以上のデータブロックの読出要求に応じて、複数のデータ領域 52bの中から、一以上のデータ領域 52b を判定する。

具体的には、読出領域判定部 12a は、ホスト装置 3 から読出要求を受けた場合に、指定されたデータが A 領域 52b-1 及び B 領域 52b-2 のいずれのデータ領域 52b に含まれるかを、読み出し対象の一以上のデータブロックに対して割り当てられた複数のデータ領域の記憶装置ビットマップ 22a を参照して判定する。

【0104】

読出処理部 12b は、記憶装置ビットマップ 22a に基づいて、読出領域判定部 12a により判定された一以上のデータ領域 52bの中から、データ保存状態が有効である読出要求に係るデータブロックを読み出す。

読出処理部 12b は、例えば、読出要求に係るデータが A 領域 52b-1 及び B 領域 52b-2 のいずれか一方のデータ領域 52b にのみ含まれる場合には、当該データ領域 52b から読出要求に係るデータブロックを読み出して、ホスト装置 3 へ出力する。具体的には、読出処理部 12b は、読み出し対象の一以上のデータブロックに対して割り当てられた複数のデータ領域の前記管理情報に基づいて、読出領域判定部 12a により判定された一以上のデータ領域の中から、データ保存状態が有効な読み出し対象の一以上のデータブロックを読み出す処理を行なう。

【0105】

一例として、図 4 を参照して説明する。前提として、記憶装置ビットマップ 22a 及び記憶領域 52 には図 4 の (3) に示すデータが格納されているものとする。また、図 4 の (2) においてユーザデータ領域 21 に格納された LBA #00 ~ #07 のデータは、全てメモリ 20 から削除され、キャッシュミスの状態になっているものとする。

図 4 の (3) において、ホスト装置 3 により、LBA #00 ~ #07 のデータの読出要求が発行されると、読出領域判定部 12a は、記憶装置ビットマップ 22a を参照して、読出要求に係るデータがどのデータ領域 52b に保存されているかを判定する。図 4 の (3) に示す例では、A 領域 52b-1 のデータ保存状態が ALL "1" であるため、LBA #00 ~ #07 の各データブロックは、全て A 領域 52b-1 に保存されている。従って、読出領域判定部 12a は、読出要求に係るデータブロックを読み出すべきデータ領域 52b が A 領域 52b-1 であると判定する。

【0106】

また、読出処理部 12b は、読出領域判定部 12a により判定された A 領域 52b-1 から、読出要求に係る LBA #00 ~ #07 のデータブロックを読み出して、ホスト装置 3 へ出力する。

一方、読出要求に係るデータが A 領域 52b-1 及び B 領域 52b-2 の双方に離散して保存されている場合も考えられる。この場合、読出処理部 12b は、それぞれのデータ領域 52b から読み出したデータブロックについて、データマージ処理部 12c に対して読出要求に係るデータの生成を指示する。

【0107】

データマージ処理部 12c は、読出処理部 12b により、二以上のデータ領域 52b の

10

20

30

40

50

中から、読出要求に係るデータブロックがそれぞれ読み出される場合に、読み出された各々のデータブロックをマージし、読み出し対象のデータを生成する処理を行なう。データマージ処理部 12c によるマージ処理は、例えばメモリ 20 のユーザデータ領域 21 上で行なわれる。データマージ処理部 12c は、マージ処理を行なって読出要求に係るデータを生成すると、読出処理部 12b を介して、又は直接、ホスト装置 3 へ出力する。

【0108】

一例として、図 3、図 5、図 7、及び図 8 を参照して説明する。前提として、記憶装置ビットマップ 22a 及び記憶領域 52 には図 3 の (4)、並びに、図 5、図 7、及び図 8 の (3) に示すデータが格納されているものとする。また、図 3 の (3)、並びに、図 5、図 7、及び図 8 の (2) においてユーザデータ領域 21 に格納された LBA # 00 ~ # 07 のデータは、全てメモリ 20 から削除され、キャッシュミスの状態になっているものとする。

10

【0109】

ホスト装置 3 により、LBA # 00 ~ # 07 のデータの読出要求が発行されると、読出領域判定部 12a は、記憶装置ビットマップ 22a を参照して、読出要求に係るデータがどのデータ領域 52b に保存されているかを判定する。図 3 の (4) に示す例では、LBA # 00 ~ # 03 の各データブロックは B 領域 52b - 2 に保存され、LBA # 04 ~ # 07 の各データブロックは A 領域 52b - 1 に保存されている。図 5 の (3) に示す例では、LBA # 00、# 01、# 04、# 05 の各データブロックは B 領域 52b - 2 に保存され、LBA # 02、# 03、# 06、# 07 の各データブロックは A 領域 52b - 1 に保存されている。図 7 の (3) に示す例では、LBA # 01 ~ # 07 の各データブロックは B 領域 52b - 2 に保存され、LBA # 00 のデータブロックは A 領域 52b - 1 に保存されている。また、図 8 の (3) に示す例では、LBA # 01 ~ # 06 の各データブロックは B 領域 52b - 2 に保存され、LBA # 00、# 07 の各データブロックは A 領域 52b - 1 に保存されている。

20

【0110】

従って、読出領域判定部 12a は、A 領域 52b - 1 及び B 領域 52b - 2 のデータ保存状態がいずれも ALL "1" ではないため、読出要求に係るデータブロックを読み出すべきデータ領域 52b が A 領域 52b - 1 及び B 領域 52b - 2 であると判定する。

また、読出処理部 12b は、読出領域判定部 12a により判定された A 領域 52b - 1 及び B 領域 52b - 2 のそれぞれから、読出要求に係る LBA # 00 ~ # 07 のデータブロックを読み出して、ユーザデータ領域 21 へ格納する。そして、読出処理部 12b は、データマージ処理部 12c に対して読出要求に係るデータの作成を指示する。

30

【0111】

データマージ処理部 12c は、ユーザデータ領域 21 に格納された各データブロックをマージする。図 3 の (5) に示す例では、データマージ処理部 12c は、B 領域 52b - 2 からの LBA # 00 ~ # 03 の各データブロックと、A 領域 52b - 1 からの LBA # 04 ~ # 07 の各データブロックとをマージする。図 5 の (4) に示す例では、データマージ処理部 12c は、B 領域 52b - 2 からの LBA # 00、# 01、# 04、# 05 の各データブロックと、A 領域 52b - 1 からの LBA # 02、# 03、# 06、# 07 の各データブロックとをマージする。図 7 の (4) に示す例では、データマージ処理部 12c は、B 領域 52b - 2 からの LBA # 01 ~ # 07 の各データブロックと、A 領域 52b - 1 からの LBA # 00 のデータブロックとをマージする。また、図 8 の (4) に示す例では、データマージ処理部 12c は、B 領域 52b - 2 からの LBA # 01 ~ # 06 の各データブロックと、A 領域 52b - 1 からの LBA # 00、# 07 の各データブロックとをマージする。

40

【0112】

そして、データマージ処理部 12c は、マージにより得られた読出要求に係るデータをホスト装置 3 へ出力する。

なお、ここまで、ホスト装置 3 からの読出要求が第 2 サイズである場合について説明し

50

たが、読出要求が第 1 サイズの単位で発行される場合でも、読出制御部 12 は上述したものと同様の制御により対応可能である。

【0113】

〔1-4〕情報処理システムの動作例

次に、上述の如く構成された情報処理システム 1（コントローラモジュール 4）の動作例を、図 10 及び図 11 を参照して説明する。図 10 は、図 1 に示す書込制御部 11 による動作例を説明するフローチャートであり、図 11 は、図 1 に示す読出制御部 12 による動作例を説明するフローチャートである。

【0114】

〔1-4-1〕書込制御部の動作例

はじめに、図 10 を参照して書込制御部 11 の動作を説明する。

なお、前提として、書込制御部 11 がホスト装置 3 から書込コマンドを受けた状態であるとする。

図 10 に示すように、書込領域判定部 11a により、ホスト装置 3 からの書込コマンドに応じて、A 領域 52b-1 の記憶装置ビットマップ 22a が ALL “1” 又は ALL “0” であるか否かが判定される（ステップ S1）。ALL “1” 又は ALL “0” であると判定された場合（ステップ S1 の Yes ルート）、さらに、書込領域判定部 11a により、A 領域 52b-1 の記憶装置ビットマップ 22a が ALL “1” であるか否かが判定される（ステップ S2）。ALL “1” ではない、つまり ALL “0” であると判定された場合（ステップ S2 の No ルート）、書込領域判定部 11a により、書き込み対象のデータ領域 52b が A 領域 52b-1 であると判定される。そして、書込処理部 11b により、書込要求に係るユーザデータが A 領域 52b-1 に書き込まれ（ステップ S3）、処理が終了する。

【0115】

一方、ステップ S2 において、A 領域 52b-1 の記憶装置ビットマップ 22a が ALL “1” であると判定された場合（ステップ S2 の Yes ルート）、書込領域判定部 11a により、書き込み対象のデータ領域 52b が B 領域 52b-2 であると判定される。そして、書込処理部 11b により、書込要求に係るユーザデータが B 領域 52b-2 に書き込まれ（ステップ S4）、処理が終了する。

【0116】

また、ステップ S1 において、A 領域 52b-1 の記憶装置ビットマップ 22a が ALL “1” 及び ALL “0” のいずれでもない判定された場合（ステップ S1 の No ルート）、書込領域判定部 11a により、次の判定処理が行なわれる。すなわち、書込領域判定部 11a により、メモリビットマップ 22b と A 領域 52b-1 又は B 領域 52b-2 の記憶装置ビットマップ 22a との論理和が ALL “1” であるか否かが判定される（ステップ S5）。論理和が ALL “1” であると判定された場合（ステップ S5 の Yes ルート）、書込領域判定部 11a により、メモリビットマップ 22b と A 領域 52b-1 の記憶装置ビットマップ 22a との論理和が ALL “1” であるか否かが判定される（ステップ S6）。

【0117】

一方、メモリビットマップ 22b と A 領域 52b-1 の記憶装置ビットマップ 22a との論理和が ALL “1” ではないと判定された場合（ステップ S6 の No ルート）、ステップ S3 に移行する。また、メモリビットマップ 22b と A 領域 52b-1 の記憶装置ビットマップ 22a との論理和が ALL “1” であると判定された場合（ステップ S6 の Yes ルート）、ステップ S4 に移行する。

【0118】

また、ステップ S5 において、論理和が ALL “1” ではないと判定された場合（ステップ S5 の No ルート）、書込領域判定部 11a により、バックエンド負荷が閾値以上であるか否かが判定される（ステップ S7）。バックエンド負荷が閾値以上であると判定された場合（ステップ S7 の Yes ルート）、書込領域判定部 11a により、マージ量が少

10

20

30

40

50

ないのはA領域52b-1であるか否かが判定される(ステップS8)。マージ量が少ないのはA領域52b-1であると判定された場合(ステップS8のYesルート)、ステップS3に移行する。一方、マージ量が少ないのはA領域52b-1ではないと判定された場合(ステップS8のNoルート)、ステップS4に移行する。

【0119】

さらに、ステップS7において、バックエンド負荷が閾値未満であると判定された場合(ステップS7のNoルート)、書込領域判定部11aにより、マージ量が多いのはA領域52b-1であるか否かが判定される(ステップS9)。マージ量が多いのはA領域52b-1であると判定された場合(ステップS9のYesルート)、ステップS3に移行する。一方、マージ量が多いのはA領域52b-1ではないと判定された場合(ステップS9のNoルート)、ステップS4に移行する。

10

【0120】

〔1-4-2〕読出制御部の動作例

次に、図11を参照して読出制御部12の動作を説明する。

なお、前提として、読出制御部12がホスト装置3から読出要求を受けた状態であるとする。

図11に示すように、読出領域判定部12aにより、読出要求に係る読み出し対象LBAの範囲のA領域52b-1の記憶装置ビットマップ22aが、ALL“1”であるか否かが判定される(ステップS11)。ALL“1”であると判定された場合(ステップS11のYesルート)、読出処理部12bにより、A領域52b-1から読出データの読み出しが行なわれ(ステップS12)、処理が終了する。

20

【0121】

一方、ステップS11において、ALL“1”ではないと判定された場合(ステップS11のNoルート)、読出領域判定部12aにより、当該A領域52b-1の記憶装置ビットマップ22aが、ALL“0”であるか否かが判定される(ステップS13)。ALL“0”であると判定された場合(ステップS13のYesルート)、読出処理部12bにより、B領域52b-2から読出データの読み出しが行なわれ(ステップS14)、処理が終了する。

【0122】

また、ステップS13において、ALL“0”ではないと判定された場合(ステップS13のNoルート)、データマージ処理部12cによる処理が行なわれる(ステップS15)。すなわち、データマージ処理部12cにより、A領域52b-1の記憶装置ビットマップ22aのうち、ビット“1”のデータブロックがA領域52b-1から、ビット“0”のデータブロックがB領域52b-2から、それぞれ読み出されてマージされる。そして、処理が終了する。

30

【0123】

〔1-5〕まとめ

上述のように、本実施形態に係るコントローラモジュール4によれば、書き込み対象の第1サイズのデータブロックを含む第2サイズの書込用データブロック(書込用データ)が、複数のデータ領域52bのうちの一のデータ領域52bへ書き込まれる。具体的には、コントローラモジュール4が、LBAのブロックサイズよりも大きい複数のセクタのうち、RMWが生じない一のセクタに、LBAで指定された対象データを含むセクタサイズのデータを書き込む。従って、HDD5等の記憶装置において書き込み対象のデータブロックの書き込みに伴う第2サイズのデータブロックの読み出し(RMW)の発生を抑制でき、HDD5の性能低下を軽減することができる。

40

【0124】

4KセクタHDD(512E)においては、RMWが頻発する可能性もあり、従来の512バイト/セクタのHDDと比較して、性能低下は顕著である。また、近年、HDDの容量の増加に伴いHDDへのアクセス頻度も増加している。HDDの容量が増加しても、ディスクの数は変わらないため、HDDはビジー状態に陥りやすくなり、性能低下が懸念

50

される。

【 0 1 2 5 】

本実施形態においては、4 KセクタHDD (5 1 2 E) におけるRMWの発生に伴う性能低下、及び、HDDが大容量化することによる性能低下に着眼し、HDD等の記憶容量の一部を用いて、RMWの発生の抑制を実現するものである。これにより、本実施形態においては、可能な限りRMWの発生を抑制してHDD 5 への書き込みを行なうことができるため、従来の5 1 2 バイト/セクタのHDDと比べても同等な書込性能を得ることができる。

【 0 1 2 6 】

なお、RMWの発生を抑制するために、HDDのデータ領域内のデータをキャッシュメモリ上に記憶しておくことも考えられるが、HDDへの書込サイズ単位でキャッシュメモリが占有されてしまい、キャッシュメモリの使用効率が悪い。また、キャッシュメモリの使用効率の悪さをキャッシュメモリの増設により補おうとしても、HDDのように大容量を容易に拡張することは困難である。

【 0 1 2 7 】

本実施形態では、HDDの記憶容量を二倍使用することになるが、近年大容量で安価なHDDが多く普及するようになり、比較的安価にHDDを増設することができる。従って、本実施形態に係るコントローラモジュール4は、HDD等の記憶装置の容量を犠牲にしても従来の5 1 2 バイト/セクタと同等なライト性能を維持することを期待する場合に、特に有効である。

【 0 1 2 8 】

また、本実施形態に係るコントローラモジュール4によれば、書き込み対象のデータブロックに対して割り当てられた複数の記憶領域のデータ保存状態が、該データブロックが書き込まれたデータ領域5 2 bに応じた値を示すように、記憶装置ビットマップ2 2 aが書き換えられる。従って、複数のデータ領域5 2 b間で書き込み対象のデータブロックの整合性を保つことができ、書込領域判定部1 1 aは、RMWが発生しない一のデータ領域5 2 bの判定を確実に行なうことができる。

【 0 1 2 9 】

また、書込領域判定部1 1 aにより、図1 0のステップS 1、S 2、S 5、及びS 6において、記憶装置ビットマップ2 2 a及びメモリビットマップ2 2 bに基づき、書き込み対象のデータブロックを書き込むべき一のデータ領域5 2 bが判定される。従って、複数のデータ領域5 2 b間で書き込み対象のデータブロックの整合性を保つことができ、書込領域判定部1 1 aは、RMWが発生しない一のデータ領域5 2 bの判定を確実に行なうことができる。

【 0 1 3 0 】

さらに、書込領域判定部1 1 aによれば、図1 0のステップS 7 ~ S 9において、書き込み対象のデータブロックを含む所定のサイズの書込用データブロックが、所定の条件に応じた最適なデータ領域5 2 bへ書き込まれる。従って、今回又は次の書き込みにおけるRMWの影響を軽減することができる。

また、記憶装置ビットマップ2 2 aは、A領域5 2 b - 1及びB領域5 2 b - 2のいずれか一方のデータ領域5 2 bに含まれるデータブロックごとに該データブロックのデータ保存状態が有効か否かを示すビットマップ情報である。従って、一のデータ領域5 2 bに係るビットマップ情報により、二つのデータ領域5 2 bにおけるデータ保存状態を管理でき、記憶装置ビットマップ2 2 aを記憶するリソースを削減することができる。

【 0 1 3 1 】

さらに、記憶装置ビットマップ管理部1 3により、書込用データブロックが一のデータ領域5 2 bへ書き込まれると、書込用データブロックに含まれる第1サイズのデータブロックについて、該データブロックに対して割り当てられた複数の記憶領域のデータ保存状態が更新される。従って、書込領域判定部1 1 aは、最新のデータ保存状態に基づき、RMWが発生しない一のデータ領域5 2 bの判定を確実に行なうことができる。

【 0 1 3 2 】

また、読出領域判定部 1 2 a により、データの読出要求に応じて、複数のデータ領域 5 2 b の中から、読み出し対象のデータを形成する一以上のデータブロックを読み出す一以上のデータ領域 5 2 b が判定される。そして、読出処理部 1 2 b によって、読出領域判定部 1 2 a により判定された一以上のデータ領域の中から、データ保存状態が有効な前記読み出し対象の一以上のデータブロックが読み出される。従って、読出制御部 1 2 は、読出要求に係る一以上のデータブロックを一以上のデータ領域 5 2 b の中から確実に読み出すことができる。

【 0 1 3 3 】

さらに、データマージ処理部 1 2 c により、複数のデータ領域 5 2 b のうちの二以上のデータ領域の中から、読み出された各々のデータブロックがマージされ、読み出し対象のデータが生成される。従って、読出要求に係るデータが複数のデータ領域 5 2 b に離散していても、マージ処理により確実に読み出すことができる。

また、書込処理部 1 1 b により、ユーザデータ領域 2 1 に保持された書き込み対象のデータブロックと、ユーザデータ領域 2 1 に保持された不要データブロックとを用いて書込用データブロックが生成される。そして、書込処理部 1 1 b により、生成された書込用データブロックが、一のデータ領域 5 2 b へ書き込まれる。従って、書込処理部 1 1 b は、書き込み対象のデータブロックに対してダミーデータ等を付加するといった処理が不要となり、書込用データブロックを容易に生成することができる。

【 0 1 3 4 】

〔 2 〕変形例

上述した一実施形態において、コントローラモジュール 4 の CPU 1 0 は、図 1 2 に示すようにさらに最適化部 1 6 を備えてもよく、書込制御部 1 1 は、図 1 に示す書込処理部 1 1 b に対して追加の処理を行なう書込処理部 1 1 b を備えてもよい。

図 1 2 は、一実施形態に係る情報処理システム 1 の構成の変形例を示す図である。なお、図 1 2 において、図 1 に示す符号と同一の符号は、図 1 に示す構成と同一又は略同一のため、重複した説明は省略する。また、図 1 2 においては、読出制御部 1 2、記憶装置ビットマップ管理部 1 3、及びビットマップ領域 2 2 の詳細な図示を省略している。

【 0 1 3 5 】

図 1 2 に示す CPU 1 0 は、図 1 に示す CPU 1 0 と比べて、書込制御部 1 1 が書込処理部 1 1 b に代えて書込処理部 1 1 b を備える点、及び、新たに最適化部 1 6 を備える点異なる。

最適化部 1 6 は、複数のデータ領域 5 2 b 間で、各々のデータ領域 5 2 b が格納する複数のデータブロックを移動させるものである。具体的には、最適化部 1 6 は、所定のタイミングで、記憶装置ビットマップ管理部 1 3 が管理する記憶装置ビットマップ 2 2 a に基づいて、複数のデータ領域 5 2 b の各々からデータ保存状態が有効な第 1 サイズのデータブロックを全て読み出す。そして、最適化部 1 6 は、読み出したデータブロックを含む第 2 サイズの書込用データブロック（書込用データ）を生成して、生成した書込用データブロックを、複数のデータ領域 5 2 b のいずれか一のデータ領域 5 2 b に書き込む。

【 0 1 3 6 】

つまり、最適化部 1 6 は、例えば図 4 の（ 1 ）のように、複数のデータ領域 5 2 b 間で有効なデータブロックが離散している場合に、それぞれのデータ領域 5 2 b から有効なデータブロックを例えばユーザデータ領域 2 1 に読み出す。そして、最適化部 1 6 は、読み出した複数のデータブロックを 4 0 9 6 バイトの書込用データブロックにマージして、複数のデータ領域 5 2 b のうちの一のデータ領域 5 2 b に書き込む（例えば図 4 の（ 3 ）セクタ # 0 - 0 参照）。

【 0 1 3 7 】

なお、所定のタイミングの例としては、以下の（ i ）～（ ii ）のうちの一つ又はこれらの組み合わせが挙げられる。

（ i ）定期的なタイミング

(ii) HDD 5 の負荷等のバックエンド負荷が閾値未満になったタイミング

このように、最適化部 16 によれば、データ保存が有効であるデータブロックを一つのデータ領域 52b に偏らせることができ、次の書き込みの際に RMW が発生しない一のデータ領域 52b が判定される確率を高めることができる。つまり、次の書き込みの際には、書込制御部 11 は、図 10 のステップ S1 及び S2 の判定を経て、RMW を発生させずに書き込み対象のデータブロックを確実に A 領域 52b-1 又は B 領域 52b-2 に書き込むことができる。従って、HDD 5 の性能低下を確実に防止することができる。

【0138】

書込処理部 11b は、図 1 に示す書込処理部 11b の機能に加えて、データマージ処理部 12c により生成されたデータを、該データを形成する一以上のデータブロックに対して割り当てられた複数のデータ領域 52b のうちの一のデータ領域 52b に書き込む機能を有する。すなわち、書込処理部 11b は、読出制御部 12 により複数のデータ領域 52b から読み出されてユーザデータ領域 21 においてマージされたデータを、一のデータ領域 52b に書き込むのである。

【0139】

このように、書込処理部 11b によっても、最適化部 16 と同様の効果を奏することができる。

なお、CPU 10 は、最適化部 16 及び書込処理部 11b の機能のうちの少なくとも一方を備えることとしてもよい。また、最適化部 16 及び書込処理部 11b の機能は、例えば書込制御部 11 及び読出制御部 12 の少なくとも一方に備えさせてもよい。

【0140】

このように、一実施形態の変形例に係る情報処理システム 1 によれば、上述した一実施形態の情報処理システム 1 と同様の効果を奏することができるほか、上述した最適化部 16 及び書込処理部 11b による効果を奏することもできる。

〔3〕ハードウェア構成例

図 13 は、図 1 又は図 12 に示すコントローラモジュール 4 のハードウェア構成例を示す図である。

【0141】

図 13 に示すように、ストレージ装置 2 は、図 1 又は図 12 において既述の CPU 10、メモリ 20、SSD 30 を備えるとともに、入出力装置 41、記録媒体 42、及び、読取装置 43 を備える。

入出力装置 41 は、例えばマウスやキーボード等の入力装置及びディスプレイやプリンタ等の出力装置の少なくとも一方を含むものであり、ユーザ操作等による動作命令を受け付ける一方、コントローラモジュール 4 による処理結果を出力する。

【0142】

記録媒体 42 は、フラッシュメモリや ROM 等の記憶装置であり、種々のデータやプログラムを記録する。

読取装置 43 は、光ディスクや USB メモリ等のコンピュータ読取可能な記録媒体 43a に記録されたデータやプログラムを読み出す装置である。

記録媒体 42 又は / 及び 43a には、一実施形態及び変形例に係る処理部の機能を実現する制御プログラムが格納されてもよい。すなわち、CPU 10 は、記録媒体 42、又は、読取装置 43 を介して記録媒体 43a から入力された制御プログラムを、メモリ 20 等の記憶装置に展開して実行することにより、処理部の機能を実現する。

【0143】

なお、上述した各ハードウェアは、互いにバスで通信可能に接続される。

〔4〕その他

以上、本発明の好ましい実施形態について詳述したが、本発明は、係る特定の実施形態に限定されるものではなく、本発明の趣旨を逸脱しない範囲内において、種々の変形、変更して実施することができる。

【0144】

例えば、一実施形態及び変形例において、データ領域 5 2 b の組は二つであるものとして説明したが、これに限定されるものではない。コントローラモジュール 4 は、三つ以上のデータ領域 5 2 b を H D D 5 の記憶領域 5 2 に管理し、三つ以上のデータ領域 5 2 b で排他的にデータブロックが保存されるように制御してもよい。

また、書込処理部 1 1 b は、ユーザデータを H D D 5 ヘライトバックするものとして説明したが、これに限定されるものではない。例えば、書込処理部 1 1 b がライトスルーを行なう場合であっても、上述した処理部の構成を適用することができる。

【 0 1 4 5 】

さらに、A 領域 5 2 b - 1 及び B 領域 5 2 b - 2 は、一つの H D D 5 の記憶領域 5 2 に割り当てられるものとして説明したが、これに限定されるものではない。A 領域 5 2 b - 1 及び B 領域 5 2 b - 2 は、複数の H D D 5 に跨って、それぞれ所定の単位で複数のセクタがまとまって割り当てられてもよい。

10

また、一実施形態及び変形例において、コントローラモジュール 4 は、ストレージ装置 2 に備えられるものとして説明したが、これに限定されるものではない。例えば、コントローラモジュール 4 は、個々の H D D 5 内に備えられてもよいし、ホスト装置 3 に備えられてもよい。

【 0 1 4 6 】

さらに、図 1 3 を参照して、一実施形態及び変形例に係るコントローラモジュール 4 は、入出力装置 4 1、記録媒体 4 2、及び読取装置 4 3 を備えるものとして説明したが、これに限定されるものではない。例えば、入出力装置 4 1、記録媒体 4 2、及び読取装置 4 3 の少なくとも一つは、ホスト装置 3 に備えられてもよい。制御プログラムはホスト装置 3 からコントローラモジュール 4 に転送され、コントローラモジュール 4 の C P U 1 0 は、ホスト装置 3 から転送された制御プログラムをメモリ 2 0 に展開して実行してもよい。また、ホスト装置 3 及びコントローラモジュール 4 の少なくとも一方は、有線又は無線を介してインターネット等のネットワークに接続されてもよい。C P U 1 0 は、制御プログラムを提供する他の装置からネットワークを介して制御プログラムを受信し、受信した制御プログラムをメモリ 2 0 に展開して実行してもよい。

20

【 0 1 4 7 】

また、一実施形態及び変形例に係るコントローラモジュール 4 は、複数の 4 K セクタ H D D (5 1 2 E) や、複数の 4 K セクタ H D D (5 1 2 E) を用いた種々の Redundant Array of Inexpensive Disks (R A I D) 構成のストレージに対しても同様の制御を行なうことができる。

30

なお、一実施形態及び変形例に係るストレージ装置 2 において、チャネルアダプタ 6 及びデバイスアダプタ 7 の少なくとも一方は、コントローラモジュール 4 に備えられてもよい。

【 0 1 4 8 】

また、一実施形態のコントローラモジュール 4 の各種機能の全部もしくは一部は、コンピュータ (C P U , 情報処理装置 , 各種端末を含む) が所定のプログラムを実行することによって実現されてもよい。

そのプログラムは、例えばフレキシブルディスク , C D (C D - R O M , C D - R , C D - R W など) , D V D (D V D - R O M , D V D - R A M , D V D - R , D V D - R W , D V D + R , D V D + R W など) , ブルーレイディスク等のコンピュータ読取可能な記録媒体に記録された形態で提供される。この場合、コンピュータはその記録媒体からプログラムを読み取って内部記憶装置または外部記憶装置に転送し格納して用いる。

40

【 0 1 4 9 】

ここで、コンピュータとは、ハードウェアと O S (オペレーティングシステム) とを含む概念であり、O S の制御の下で動作するハードウェアを意味している。また、O S が不要でアプリケーションプログラム単独でハードウェアを動作させるような場合には、そのハードウェア自体がコンピュータに相当する。ハードウェアは、少なくとも、C P U 等のマイクロプロセッサと、記録媒体に記録されたコンピュータプログラムを読み取る手段と

50

をそなえている。上記プログラムは、上述のようなコンピュータに、一実施形態のコントローラモジュール（制御装置，制御部）4の各種機能を実現させるプログラムコードを含んでいる。また、その機能の一部は、アプリケーションプログラムではなくOSによって実現されてもよい。

【0150】

〔5〕付記

以上の実施形態及び変形例に関し、更に以下の付記を開示する。

（付記1）

上位装置に対して第1サイズ of データブロック単位で入出力を行ない、記憶装置の各データ領域に対して前記第1サイズと異なる第2サイズ単位でデータの入出力を行なう制御装置であって、

10

前記データ領域を複数の第1サイズのデータブロックの記憶領域に割り当て、複数の前記データ領域に同じ第1サイズのデータブロックを割り当てる処理を行なう割当処理部と、

各データ領域における前記複数の第1サイズのデータブロックの記憶領域のデータ保存状態を示す第1管理情報を管理する管理部と、

書き込み対象のデータブロックに対して割り当てられた複数のデータ領域の前記第1管理情報に基づいて、該複数のデータ領域の中から一のデータ領域を判定する書込領域判定部と、

前記書き込み対象のデータブロックのデータを含む前記第2サイズの書込用データを生成し、生成した前記書込用データを、前記書込領域判定部により判定された前記一のデータ領域に書き込む処理を行なう書込処理部と、

20

を備えることを特徴とする、制御装置。

【0151】

（付記2）

前記管理部は、

前記書き込み対象のデータブロックに対して割り当てられた複数の記憶領域のデータ保存状態が、該複数の記憶領域のうち、一の記憶領域は前記データブロックのデータが書き込まれると有効を示し、他の全ての記憶領域は無効を示すように、前記第1管理情報を書き換える

30

ことを特徴とする、付記1記載の制御装置。

【0152】

（付記3）

前記管理部は、

前記各データ領域に割り当てられた前記複数のデータブロックの記憶領域について、データブロックごとに、前記複数のデータ領域のうちのいずれか一のデータ領域内で該データブロックに対して割り当てられた記憶領域が、他の全てのデータ領域内で該データブロックに対して割り当てられた記憶領域との間で排他的に有効を示すように、前記データ保存状態を管理し、

40

前記書込領域判定部は、

前記書き込み対象のデータブロックに対して割り当てられた複数のデータ領域の前記第1管理情報に基づいて、前記複数のデータ領域のうちの一つである第1のデータ領域における全てのデータブロックのデータ保存状態が無効であると判断した場合に、前記第1のデータ領域を、前記一のデータ領域と判定する

ことを特徴とする、付記2記載の制御装置。

【0153】

（付記4）

前記書込領域判定部は、

前記書き込み対象のデータブロックに対して割り当てられた複数のデータ領域の前記第1管理情報と、前記書き込み対象のデータブロックに対して割り当てられた前記複数のデ

50

ータ領域における前記書き込み対象のデータブロックが保存される領域を示す第2管理情報とに基づき、前記第1のデータ領域における前記データ保存状態が有効な全てのデータブロックの領域が、前記第1のデータ領域において前記書き込み対象のデータブロックが保存される全ての領域に含まれると判断した場合に、前記第1のデータ領域を、前記一のデータ領域と判定する

ことを特徴とする、付記3記載の制御装置。

【0154】

(付記5)

前記書込領域判定部は、

前記第1及び第2管理情報に基づく判断によっても、前記複数のデータ領域の中から前記一のデータ領域が判定できない場合に、所定の条件に応じて、前記複数のデータ領域からいずれか一のデータ領域を選択し、

前記書込処理部は、

選択した前記一のデータ領域について、前記書き込み対象のデータブロックに対して割り当てられた複数のデータ領域の前記第1管理情報に基づいて、前記書き込み対象のデータブロックを含む所定のサイズの書込用データを生成し、生成した前記書込用データを、前記書込領域判定部により選択された前記一のデータ領域へ書き込む

ことを特徴とする、付記4記載の制御装置。

【0155】

(付記6)

前記複数のデータ領域は、第1のデータ領域及び第2のデータ領域の二つのデータ領域であり、

前記第1管理情報は、前記第1及び第2のデータ領域のいずれか一方のデータ領域に含まれる前記データブロックごとに該データブロックのデータ保存状態が有効か否かを示すビットマップ情報であり、

前記管理部は、

前記ビットマップ情報に基づいて、前記第1及び第2のデータ領域におけるデータ保存状態を管理する

ことを特徴とする、付記3～5のいずれか1項記載の制御装置。

【0156】

(付記7)

所定のタイミングで、前記管理部が管理する前記データ保存状態に基づいて、前記複数のデータ領域の各々から前記データ保存状態が有効な前記第1サイズのデータブロックを全て読み出し、読み出した前記データブロックを含む前記第2サイズの書込用データを生成して、生成した前記書込用データを、前記複数のデータ領域のいずれか一のデータ領域に書き込む最適化部、

をさらに備えることを特徴とする、付記2～6のいずれか1項記載の制御装置。

【0157】

(付記8)

前記管理部は、

前記書込用データが前記一のデータ領域へ書き込まれると、前記書込用データに含まれる前記第1サイズのデータブロックについて、該データブロックに対して割り当てられた複数の記憶領域のデータ保存状態を更新する

ことを特徴とする、付記2～7のいずれか1項記載の制御装置。

【0158】

(付記9)

読み出し対象のデータを形成する一以上のデータブロックに対して割り当てられた複数のデータ領域の前記第1管理情報に基づいて、前記複数のデータ領域の中から、一以上のデータ領域を判定する読出領域判定部と、

前記読み出し対象の一以上のデータブロックに対して割り当てられた複数のデータ領域

10

20

30

40

50

の前記第 1 管理情報に基づいて、前記読出領域判定部により判定された前記一以上のデータ領域の中から、前記データ保存状態が有効な前記読み出し対象の一以上のデータブロックを読み出す処理を行なう読出処理部と、
をさらに備えることを特徴とする、付記 2 ～ 8 のいずれか 1 項記載の制御装置。

【0159】

(付記 10)

前記読出処理部により、前記複数のデータ領域のうちの二以上のデータ領域の中からデータブロックがそれぞれ読み出されると、読み出された各々のデータブロックをマージして、前記読み出し対象のデータを生成する処理を行なうデータマージ処理部、
をさらに備えることを特徴とする、付記 9 記載の制御装置。

10

【0160】

(付記 11)

前記書込処理部は、
前記データマージ処理部により生成された前記データを、該データを形成する前記一以上のデータブロックに対して割り当てられた複数のデータ領域のうちの一のデータ領域に書き込む
ことを特徴とする、付記 10 記載の制御装置。

【0161】

(付記 12)

前記複数のデータ領域のいずれか一つに書き込む前記書き込み対象のデータブロックを保持する前記第 2 サイズ以上の保持領域を有する保持部、
をさらに備え、

20

前記書込処理部は、

前記保持領域に保持された前記書き込み対象のデータブロックと、前記保持領域に保持された不要データブロックとを用いて前記書込用データを生成し、生成した前記書込用データを、前記書込領域判定部により判定された前記一のデータ領域へ書き込む
ことを特徴とする、付記 1 ～ 11 のいずれか 1 項記載の制御装置。

【0162】

(付記 13)

データを記憶する記憶装置と、
上位装置に対して第 1 サイズのデータブロック単位で入出力を行ない、前記記憶装置の各データ領域に対して前記第 1 サイズと異なる第 2 サイズ単位でデータの入出力を行なう制御装置と、
を備え、

30

前記制御装置は、

前記データ領域を複数の第 1 サイズのデータブロックの記憶領域に割り当て、複数の前記データ領域に同じ第 1 サイズのデータブロックを割り当てる処理を行なう割当処理部と、

各データ領域における前記複数の第 1 サイズのデータブロックの記憶領域のデータ保存状態を示す第 1 管理情報を管理する管理部と、

40

書き込み対象のデータブロックに対して割り当てられた複数のデータ領域の前記第 1 管理情報に基づいて、該複数のデータ領域の中から一のデータ領域を判定する書込領域判定部と、

前記書き込み対象のデータブロックのデータを含む前記第 2 サイズの書込用データを生成し、生成した前記書込用データを、前記書込領域判定部により判定された前記一のデータ領域に書き込む処理を行なう書込処理部と、
を備えることを特徴とする、ストレージ装置。

【0163】

(付記 14)

前記管理部は、

50

前記書き込み対象のデータブロックに対して割り当てられた複数の記憶領域のデータ保存状態が、該複数の記憶領域のうち、一の記憶領域は前記データブロックのデータが書き込まれると有効を示し、他の全ての記憶領域は無効を示すように、前記第 1 管理情報を書き換える

ことを特徴とする、付記 13 記載のストレージ装置。

【0164】

(付記 15)

前記管理部は、

前記各データ領域に割り当てられた前記複数のデータブロックの記憶領域について、データブロックごとに、前記複数のデータ領域のうちのいずれか一のデータ領域内で該データブロックに対して割り当てられた記憶領域が、他の全てのデータ領域内で該データブロックに対して割り当てられた記憶領域との間で排他的に有効を示すように、前記データ保存状態を管理し、

10

前記書込領域判定部は、

前記書き込み対象のデータブロックに対して割り当てられた複数のデータ領域の前記第 1 管理情報に基づいて、前記複数のデータ領域のうちの一つである第 1 のデータ領域における全てのデータブロックのデータ保存状態が無効であると判断した場合に、前記第 1 のデータ領域を、前記一のデータ領域と判定する

ことを特徴とする、付記 14 記載のストレージ装置。

【0165】

20

(付記 16)

前記書込領域判定部は、

前記書き込み対象のデータブロックに対して割り当てられた複数のデータ領域の前記第 1 管理情報と、前記書き込み対象のデータブロックに対して割り当てられた前記複数のデータ領域における前記書き込み対象のデータブロックが保存される領域を示す第 2 管理情報とに基づき、前記第 1 のデータ領域における前記データ保存状態が有効な全てのデータブロックの領域が、前記第 1 のデータ領域において前記書き込み対象のデータブロックが保存される全ての領域に含まれると判断した場合に、前記第 1 のデータ領域を、前記一のデータ領域と判定する

ことを特徴とする、付記 15 記載のストレージ装置。

30

【0166】

(付記 17)

前記制御装置は、

所定のタイミングで、前記管理部が管理する前記データ保存状態に基づいて、前記複数のデータ領域の各々から前記データ保存状態が有効な前記第 1 サイズのデータブロックを全て読み出し、読み出した前記データブロックを含む前記第 2 サイズの書込用データを生成して、生成した前記書込用データを、前記複数のデータ領域のいずれか一のデータ領域に書き込む最適化部、

をさらに備えることを特徴とする、付記 13 ~ 16 のいずれか 1 項記載のストレージ装置

40

【0167】

(付記 18)

前記制御装置は、

読み出し対象のデータを形成する一以上のデータブロックに対して割り当てられた複数のデータ領域の前記第 1 管理情報に基づいて、前記複数のデータ領域の中から、一以上のデータ領域を判定する読出領域判定部と、

前記読み出し対象の一以上のデータブロックに対して割り当てられた複数のデータ領域の前記第 1 管理情報に基づいて、前記読出領域判定部により判定された前記一以上のデータ領域の中から、前記データ保存状態が有効な前記読み出し対象の一以上のデータブロックを読み出す処理を行なう読出処理部と、

50

をさらに備えることを特徴とする、付記 13 ~ 16 のいずれか 1 項記載のストレージ装置。

【0168】

(付記 19)

前記制御装置は、

前記読出処理部により、前記複数のデータ領域のうちの二以上のデータ領域の中からデータブロックがそれぞれ読み出されると、読み出された各々のデータブロックをマージして、前記読み出し対象のデータを生成する処理を行なうデータマージ処理部、
をさらに備えることを特徴とする、付記 18 記載のストレージ装置。

【0169】

(付記 20)

上位装置に対して第 1 サイズのデータブロック単位で入出力を行ない、記憶装置の各データ領域に対して前記第 1 サイズと異なる第 2 サイズ単位でデータの入出力を行なうコンピュータに、

前記データ領域を複数の第 1 サイズのデータブロックの記憶領域に割り当て、複数の前記データ領域に同じ第 1 サイズのデータブロックを割り当て、

各データ領域における前記複数の第 1 サイズのデータブロックの記憶領域のデータ保存状態を示す管理情報を管理し、

書き込み対象のデータブロックに対して割り当てられた複数のデータ領域の前記管理情報に基づいて、該複数のデータ領域の中から一のデータ領域を判定し、

前記書き込み対象のデータブロックのデータを含む前記第 2 サイズの書込用データを生成し、生成した前記書込用データを、判定した前記一のデータ領域に書き込む、
処理を実行させることを特徴とする、制御プログラム。

【符号の説明】

【0170】

1, 100 情報処理システム

2, 200 ストレージ装置

3 ホスト装置 (上位装置)

4 コントローラモジュール (制御装置, 制御部)

5 HDD (記憶装置)

6 チャネルアダプタ

7 デバイスアダプタ

10 CPU (処理部)

11 書込制御部

11a 書込領域判定部

11b, 11b 書込処理部

12 読出制御部

12a 読出領域判定部

12b 読出処理部

12c データマージ処理部

13 記憶装置ビットマップ管理部 (管理部)

13a ビットマップ更新部

13b ビットマップ読出部

13c ビットマップ書込部

14 割当処理部

15 メモリビットマップ管理部

16 最適化部

20 メモリ (保持部)

21 ユーザデータ領域 (保持領域)

22 ビットマップ領域

10

20

30

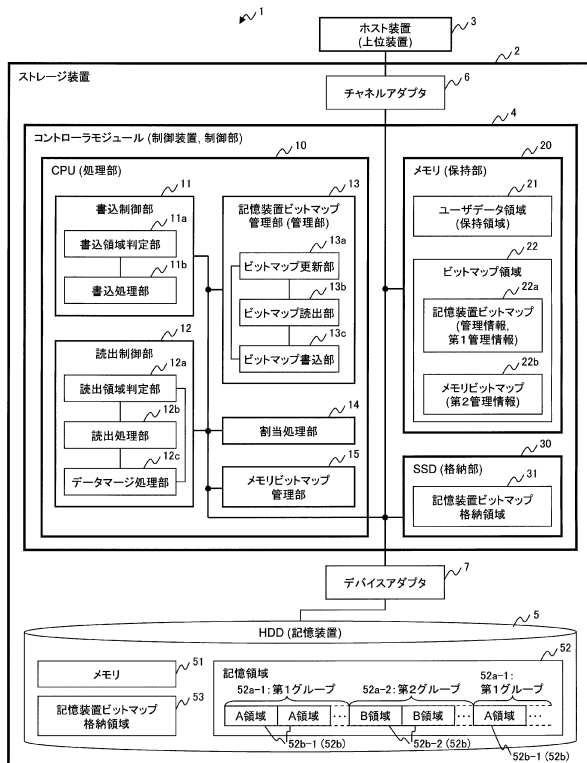
40

50

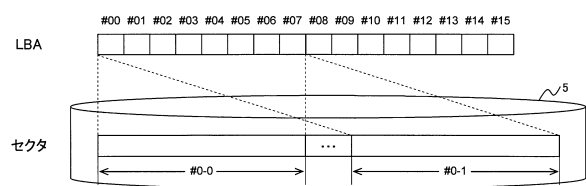
- 2 2 a 記憶装置ビットマップ (管理情報, 第1管理情報)
- 2 2 b メモリビットマップ (第2管理情報)
- 3 0 S S D (格納部)
- 3 1, 5 3 記憶装置ビットマップ格納領域
- 4 1 入出力装置
- 4 2, 4 3 a 記録媒体
- 4 3 読取装置
- 5 1, 4 1 0, 5 1 0 メモリ
- 5 2, 5 2 0 記憶領域
- 5 2 a - 1 第1グループ
- 5 2 a - 2 第2グループ
- 5 2 b, 5 2 b - 1 A領域 (データ領域)
- 5 2 b, 5 2 b - 2 B領域 (データ領域)
- 3 0 0 ホスト装置
- 4 0 0 コントローラモジュール
- 5 0 0 H D D

10

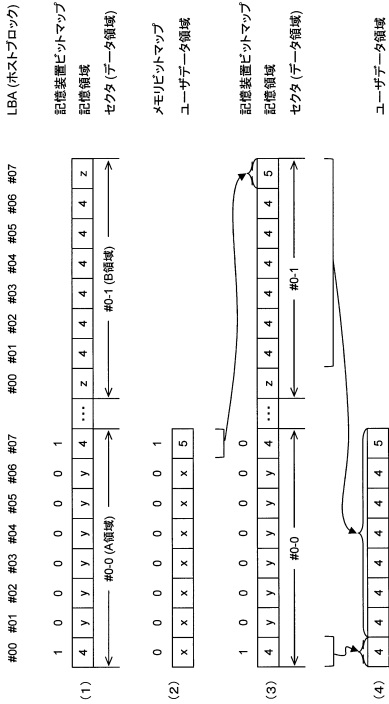
【図1】



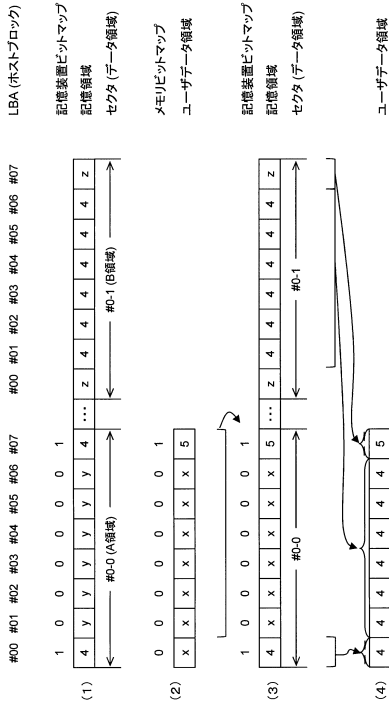
【図2】



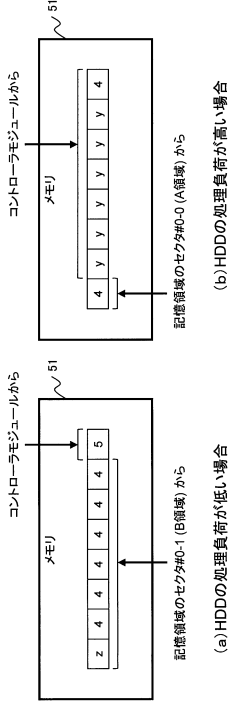
【図 7】



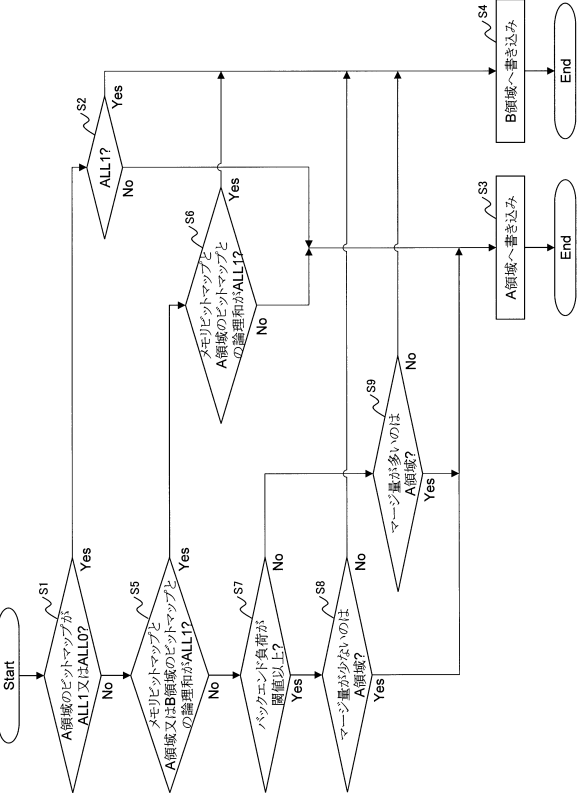
【図 8】



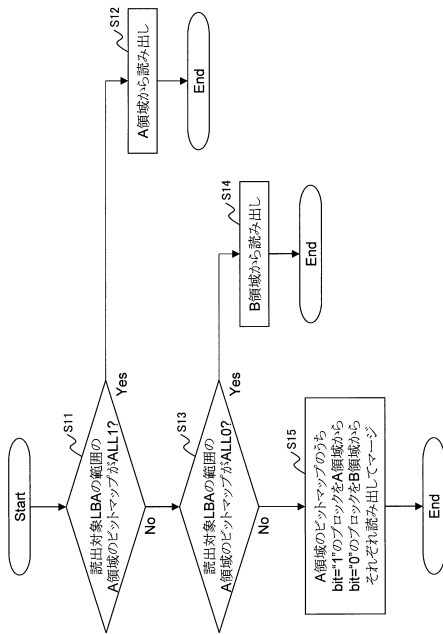
【図 9】



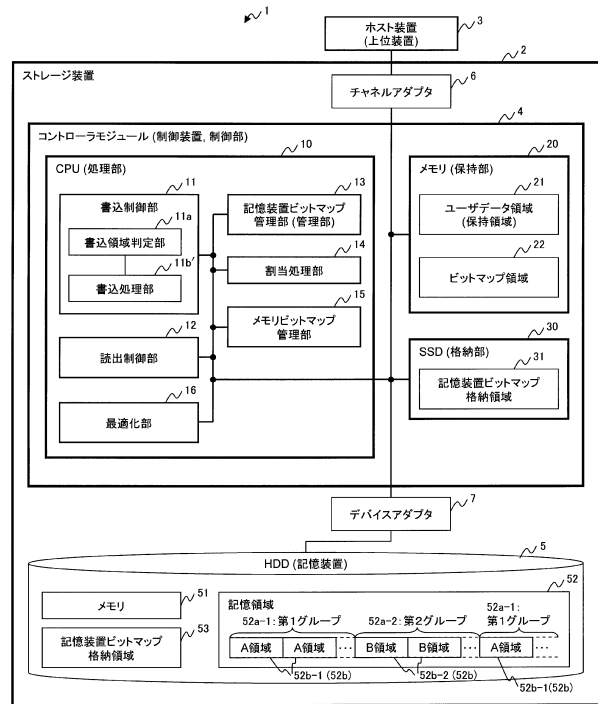
【図 10】



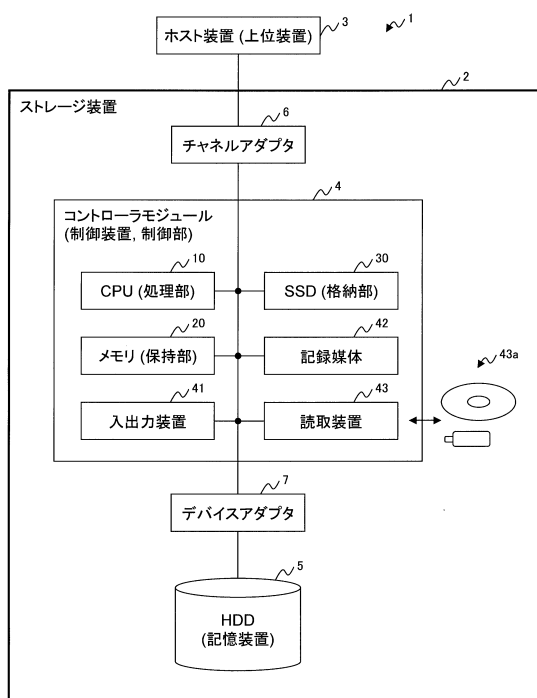
【図 1 1】



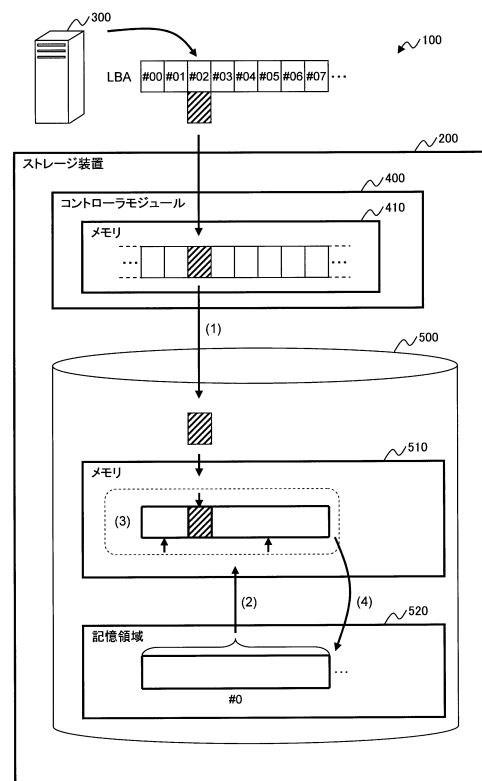
【図 1 2】



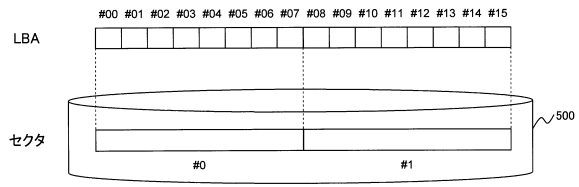
【図 1 3】



【図 1 4】



【図 15】



フロントページの続き

- (72)発明者 紺田 與志仁
神奈川県川崎市中原区上小田中4丁目1番1号 株式会社富士通コンピュータテクノロジーズ内
- (72)発明者 猪頭 惇
神奈川県川崎市中原区上小田中4丁目1番1号 株式会社富士通コンピュータテクノロジーズ内
- (72)発明者 仁村 康太郎
神奈川県川崎市中原区上小田中4丁目1番1号 株式会社富士通コンピュータテクノロジーズ内
- (72)発明者 安部 麻理恵
神奈川県川崎市中原区上小田中4丁目1番1号 株式会社富士通コンピュータテクノロジーズ内
- (72)発明者 東條 美帆子
神奈川県川崎市中原区上小田中4丁目1番1号 株式会社富士通コンピュータテクノロジーズ内
- (72)発明者 中村 政智
神奈川県川崎市中原区上小田中4丁目1番1号 株式会社富士通コンピュータテクノロジーズ内

審査官 桜井 茂行

- (56)参考文献 特開2009-134682(JP, A)
米国特許出願公開第2009/0113119(US, A1)
特開2010-250847(JP, A)
特開2007-193439(JP, A)
特開2011-222015(JP, A)
特開2010-267345(JP, A)
特開2009-032323(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06 - 3/08
G06F 12/00
G06K 1/00 - 5/04