

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6540363号
(P6540363)

(45) 発行日 令和1年7月10日(2019.7.10)

(24) 登録日 令和1年6月21日(2019.6.21)

(51) Int.Cl. F I
G06F 3/06 (2006.01)
 G06F 3/06 3 O 1 N
 G06F 3/06 3 O 1 V
 G06F 3/06 3 O 1 Z

請求項の数 7 (全 20 頁)

(21) 出願番号	特願2015-162115 (P2015-162115)	(73) 特許権者	000005223 富士通株式会社
(22) 出願日	平成27年8月19日 (2015.8.19)		神奈川県川崎市中原区上小田中4丁目1番1号
(65) 公開番号	特開2017-41076 (P2017-41076A)	(74) 代理人	100104190 弁理士 酒井 昭徳
(43) 公開日	平成29年2月23日 (2017.2.23)	(72) 発明者	山地 一史 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
審査請求日	平成30年5月14日 (2018.5.14)	審査官	田中 啓介

最終頁に続く

(54) 【発明の名称】 ストレージ制御装置、ストレージ制御方法、およびストレージ制御プログラム

(57) 【特許請求の範囲】

【請求項1】

アンアラインドとなる書き込みデータの書き込み要求に応じて前記書き込みデータに対するリードモディファイライト処理を行うストレージから受信した、前記リードモディファイライト処理に用いる記憶領域の使用状況を示す使用情報を記憶する記憶部と、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報によって前記記憶領域が使用されている状況が示される場合、自装置による当該書き込みデータに対するリードモディファイライト処理を実行する制御部と、

を有することを特徴とするストレージ制御装置。

【請求項2】

前記制御部は、

自装置がコマンドを前記ストレージに送信したことに応じた前記コマンドの実行結果とともに前記使用情報を受信する、

ことを特徴とする請求項1に記載のストレージ制御装置。

【請求項3】

前記制御部は、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報によって前記記憶領域が使用されていない状況が示される場合、当該書き込みデータの書き込み要求を前記ストレージに送信する処理を実行する、

ことを特徴とする請求項1または2に記載のストレージ制御装置。

【請求項 4】

前記制御部は、

前記ストレージへの書き込みデータが発生した際に、前記書き込みデータがアラインドかアンアラインドかを判定し、

前記書き込みデータがアンアラインドであると判定した場合、前記使用情報によって前記記憶領域が使用されている状況が示される場合に、自装置による当該書き込みデータに対するリードモディファイライト処理を実行する、

ことを特徴とする請求項 1～3 のいずれか一つに記載のストレージ制御装置。

【請求項 5】

前記制御部は、

前記書き込みデータがアンアラインドであると判定した場合、前記使用情報によって前記記憶領域が使用されていない状況が示される場合には、当該書き込みデータの書き込み要求を前記ストレージに送信する処理を実行する、

ことを特徴とする請求項 4 に記載のストレージ制御装置。

【請求項 6】

コンピュータが、

アンアラインドとなる書き込みデータの書き込み要求に応じて前記書き込みデータに対するリードモディファイライト処理を行うストレージから受信した、前記リードモディファイライト処理に用いる記憶領域の使用状況を示す使用情報を記憶部に記憶させ、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報によって前記記憶領域が使用されている状況が示される場合、自装置による当該書き込みデータに対するリードモディファイライト処理を実行する、

処理を実行することを特徴とするストレージ制御方法。

【請求項 7】

コンピュータに、

アンアラインドとなる書き込みデータの書き込み要求に応じて前記書き込みデータに対するリードモディファイライト処理を行うストレージから受信した、前記リードモディファイライト処理に用いる記憶領域の使用状況を示す使用情報を記憶部に記憶させ、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報によって前記記憶領域が使用されている状況が示される場合、自装置による当該書き込みデータに対するリードモディファイライト処理を実行する、

処理を実行させることを特徴とするストレージ制御プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ制御装置、ストレージ制御方法、およびストレージ制御プログラムに関する。

【背景技術】

【0002】

従来、所定値で割り切れるアドレス範囲へのアクセスだけが行えるストレージに対して、所定値で割り切れないアドレス範囲の書き込みデータであるアンアラインドとなる書き込みデータを書き込む際に、リードモディファイライト処理を行う技術がある。また、リードモディファイライト処理に用いる記憶領域を有するストレージがある。

【0003】

関連する先行技術として、例えば、データ記憶装置において、リードモディファイライト処理の前にバッファメモリに格納されているサブセクタの各データを結合して、ロングセクタのフォーマットに変換するものがある。また、リードモディファイライト処理の実行主体をホスト制御部とした場合の処理速度と、メディア制御部とした場合の処理速度を計測し、処理速度の速い方をリードモディファイライト処理の実行主体とする技術がある。

。

10

20

30

40

50

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2013-157068号公報

【特許文献2】特開2010-33396号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、従来技術によれば、ハードディスクドライブ（HDD：Hard Disk Drive）等のストレージの上位となる装置が、ストレージにリードモディファイライト処理を行わせると、アンアラインドとなる書き込みデータの書き込み要求に対する応答性能が低下する場合がある。例えば、リードモディファイライト処理に用いる記憶領域に空き領域がない際に、ストレージにリードモディファイライト処理を行わせると、アンアラインドとなる書き込みデータの書き込み要求に対する応答性能が低下する。

10

【0006】

1つの側面では、本発明は、アンアラインドとなる書き込みデータの書き込み要求に対する応答性能の低下を抑制できるストレージ制御装置、ストレージ制御方法、およびストレージ制御プログラムを提供することを目的とする。

【課題を解決するための手段】

【0007】

20

本発明の一側面によれば、アンアラインドとなる書き込みデータの書き込み要求に応じて書き込みデータに対するリードモディファイライト処理を行うストレージから受信した、リードモディファイライト処理に用いる記憶領域の使用状況を示す使用情報を記憶し、ストレージへのアンアラインドとなる書き込みデータが発生した際に、使用情報に基づいて、書き込みデータの書き込み要求をストレージに送信する処理、または、自装置による書き込みデータに対するリードモディファイライト処理のいずれか一方を実行するストレージ制御装置、ストレージ制御方法、およびストレージ制御プログラムが提案される。

【発明の効果】

【0008】

本発明の一態様によれば、アンアラインドとなる書き込みデータの書き込み要求に対する応答性能の低下を抑制できるという効果を奏する。

30

【図面の簡単な説明】

【0009】

【図1】図1は、本実施の形態にかかるストレージ制御装置101の動作例を示す説明図である。

【図2】図2は、情報処理システム200の一例を示す説明図である。

【図3】図3は、ディスクアレイ装置202のハードウェア構成例を示す説明図である。

【図4】図4は、管理装置201のハードウェア構成例を示す説明図である。

【図5】図5は、HDD102#0の例を示す説明図である。

【図6】図6は、CM301の機能構成例を示す説明図である。

40

【図7】図7は、ライトコマンド以外のコマンド発行時処理手順の一例を示すフローチャートである。

【図8】図8は、ライトコマンド発行時処理手順の一例を示すフローチャート（その1）である。

【図9】図9は、ライトコマンド発行時処理手順の一例を示すフローチャート（その2）である。

【図10】図10は、ライトコマンド発行時処理手順の一例を示すフローチャート（その3）である。

【図11】図11は、メディアキャッシュ状態管理テーブル611の記憶内容の一例を示す説明図である。

50

【発明を実施するための形態】**【0010】**

以下に図面を参照して、開示のストレージ制御装置、ストレージ制御方法、およびストレージ制御プログラムの実施の形態を詳細に説明する。

【0011】

図1は、本実施の形態にかかるストレージ制御装置101の動作例を示す説明図である。ストレージ制御装置101は、1つ以上のストレージを制御するコンピュータである。ストレージは、データを保存する装置である。例えば、ストレージは、例えば、HDDやSSD(Solid State Drive)等である。以下の例では、ストレージがHDDであるとする。例えば、ストレージ制御装置101は、上位となるサーバ等に、ストレージの記憶領域を提供する。

10

【0012】

また、ストレージ制御装置101は、RAID(Redundant Arrays of Inexpensive Disks)技術により、複数のHDDを組み合わせて仮想的なディスクを提供する。RAID技術には、仮想的なディスクの形成の仕方を表すRAIDレベルが存在する。ここで、本実施の形態では、RAID0~6のうちのいずれのRAIDレベルであってもよく、また、RAID1+5というように、RAIDレベルを組み合わせたレベルを対象としてもよい。

【0013】

また、近年、HDDの大容量化および1度に処理されるデータ量の増大化に伴い、物理セクタのサイズを、512バイト(Bytes)から4キロバイト(KB:Kilo-Bytes)に拡張し、HDD内のトラック上のギャップを減らすAdvanced Formatという方法がある。Advanced Formatに対応したHDDは、512バイトアクセスだけを行うホストに対応するために、HDDの内部は4KBの物理セクタでありながら、外部からは512バイトの論理セクタとしてアクセスが可能なエミュレート機能を有する。

20

【0014】

Advanced Formatに対応したHDDは、エミュレート機能により、512バイト単位でのアクセスが可能である。以下の説明では、単にHDDと記載した場合、Advanced Formatに対応したHDDであるとする。HDDは、物理セクタのサイズである4KBのパウンダリアクセスとならない書き込みにおいては、HDDの物理セクタへの書き込みは4KB単位でしか実施できない。ここで、パウンダリアクセスとは、所定値で割り切れるアドレス範囲へのアクセスである。また、所定値で割り切れるアドレス範囲のことを、「所定値+境界(パウンダリ)」と呼ぶことがある。例えば、4KBで割り切れるアドレス範囲のことを「4KB境界」と呼ぶことがある。所定値は、4KB、8KB、16KB等どのような値でもよく、本実施の形態では、所定値は、4KBとなる。

30

【0015】

また、所定値で割り切れるアドレス範囲となる条件は、アドレス範囲の先頭のアドレスと末尾のアドレスとが所定値で割り切れるアドレスとなることである。例えば、アドレス範囲の先頭のアドレスが所定値で割り切れるアドレスであり、かつ、アクセス範囲となるデータ長が、所定値で割り切れる値となるアクセスは、パウンダリアクセスとなる。

40

【0016】

以下の記載では、パウンダリアクセスとなるアクセス先のデータを、アラインド(Aligned)となるデータと称する。すなわち、アラインドとなるデータは、データの先頭アドレスと末尾のアドレスとが所定値で割り切れるアドレスとなるデータである。これに対し、パウンダリアクセスとならないアクセス先のデータを、アンアラインド(UnAligned)となるデータと称する。すなわち、アンアラインドとなるデータは、データの先頭アドレスまたは末尾のアドレスの少なくともいずれか一方のアドレスが所定値で割り切れないアドレスとなるデータである。

50

【0017】

アンアラインドとなる書き込みデータの書き込み要求を受け付けると、HDDは、以下に示すリードモディファイライト(RMW: Read Modify Write)処理を行う。まず、HDDは、アンアラインドとなる書き込みデータをメディアキャッシュ(Media Cache)と呼ばれる領域に書き込む。メディアキャッシュを含めたHDDのハードウェア構成を図5で示す。次に、HDDは、4KB境界の物理セクタのデータをHDD内のDRAM(Dynamic Random Access Memory)に読み出す(Read)。HDD内のDRAMは、任意のアドレスにアクセスすることができる記憶領域である。そして、HDDは、メディアキャッシュ上のデータについて、サーバ等の上位からの書き込み要求のあった論理セクタの書き込みデータを更新する(Mo 10
dify)。ここで、更新する際に、HDDは、HDD内のDRAMに書き込まれた4KB境界の物理セクタのデータに、アンアラインドとなる書き込みデータを上書きすることにより、更新を行う。

【0018】

次に、HDDは、更新されたメディアキャッシュ上のデータを物理セクタに書き戻す(Write)。なお、アラインドとなる書き込みデータの書き込み要求については、HDDは、RMW処理を行わず、直接磁気ディスクに書き込みデータを書き込む。

【0019】

また、RMW処理を、HDDの上位となる装置で行うこともできる。一般的に、メディアキャッシュの利用により、HDDの上位となる装置がRMW処理を行うよりもHDDが 20
RMW処理を行う方が、HDDとHDDの上位となる装置とのインタフェースを介さない分、高い書き込み性能が得られる。

【0020】

しかしながら、HDDの上位となる装置が、HDDにRMW処理を行わせると、アンアラインドとなる書き込みデータの書き込み要求に対する応答性能が低下する場合がある。例えば、メディアキャッシュの空き領域がない際に、HDDにRMW処理を行わせようとすると、メディアキャッシュに空き領域が発生するまで待機することになるため、アンアラインドとなる書き込みデータの書き込み要求に対する応答性能が低下する。ここで、メディアキャッシュの空き領域がなくなる場合は、例えば、アンアラインドとなる書き込みデータの書き込み要求がHDDに連続して送信されて、RMW処理が完了しメディアキャ 30
ッシュ上の書き込みデータを解放する処理が間に合わなくなった場合である。

【0021】

また、HDDの上位の装置により、HDDの性能テストを行い、HDDの性能テストのデータを基に、メディアキャッシュの枯渇を予測し、HDDと、HDDの上位の装置とによるRMW処理を動的に切り替えることが考えられる。しかしながら、この場合、故障したHDDを交換する運用を前提としていると、故障したHDDを交換するたびに性能テストを行うことになり、性能テスト中は交換したHDDにデータを格納できない時間が発生する。また、性能テストで書き込みを行う場合、性能テストで書き込みを行った場合は通常の書き込みと性能テストによる書き込みとなり、Write素子への負荷が性能テストによる書き込みの分大きくなるため、HDDの寿命を縮めることになる。 40

【0022】

そこで、本実施の形態では、HDDからのメディアキャッシュの使用状況に基づいて、RMW処理の主体をストレージ制御装置101とHDDとで切り替える方法について説明する。

【0023】

図1を用いて、ストレージ制御装置101の動作を説明する。ストレージ制御装置101は、HDD102#0~#Nを制御する。そして、ストレージ制御装置101は、RAID技術によるHDD102#0~#Nを組み合わせることにより得られた仮想的なディスクを、ストレージ制御装置101を利用する装置に提供する。ここで、ストレージ制御装置101は、HDD102#0~#Nにより、1つの仮想的なディスクを形成してもよ 50

いし、複数の仮想的なディスクを形成してもよい。例えば、HDD102#0～#Nにより1つの仮想的なディスクを形成する際には、HDD102#0～#Nが1つのRAIDグループに属することになる。また、HDD102#0～#Nは、Advanced Formatに対応しており、それぞれ、メディアキャッシュ103#0～#Nを有する。なお、ストレージ制御装置101は、Advanced Formatに対応するHDDと、Advanced Formatに対応していないHDDとを混在して制御してもよい。

【0024】

図1の(1)で示すように、ストレージ制御装置101は、HDD102#0～#Nから使用情報111を受信する。ストレージ制御装置101は、受信した使用情報111をストレージ制御装置101がアクセス可能な記憶部に記憶する。使用情報111は、HDD102のメディアキャッシュ103の使用状況を示す情報である。例えば、使用情報111は、メディアキャッシュ103が使用されている状況を示す識別子か、メディアキャッシュ103が使用されていない状況を示す識別子のいずれかの識別子となる。ここで、メディアキャッシュ103が使用されている状況は、メディアキャッシュ103の空き容量が所定の閾値以下である状況である。同様に、メディアキャッシュ103が使用されていない状況は、メディアキャッシュ103の空き容量が所定の閾値以上である状況である。

10

【0025】

以下、メディアキャッシュ103が使用されている状況を示す識別子を「Busy」とする。同様に、メディアキャッシュ103が使用されていない状況を示す識別子を「Free」とする。

20

【0026】

また、使用情報111を受信するタイミングは、いつでもよい。例えば、ストレージ制御装置101は、HDD102に対してコマンドを送信したことに応じて、HDD102からコマンドの実行結果を受信するとともに使用情報111を受信してもよい。ここで、コマンドは、特定の機能の実行を指示する命令である。例えば、HDD102に送信するコマンドには、あるアドレスへの書き込みデータの書き込み要求を指示する命令であるライトコマンドや、あるアドレスの読み出し要求を指示する命令であるリードコマンドや、テストユニットレディコマンド等がある。

30

【0027】

また、HDD102は、コマンドの実行結果をストレージ制御装置101に送信する際に、メディアキャッシュ103の使用状況が変化したときに限り使用情報111を送信してもよいし、常に使用情報111を送信してもよい。また、ストレージ制御装置101が使用情報111の送信要求をHDD102に送信し、送信要求を受け付けたHDD102が使用情報111をストレージ制御装置101に送信してもよい。また、ストレージ制御装置101は、メディアキャッシュ103の使用状況が変化した際にHDD102が自動的に送信した使用情報111を受信してもよい。以下の例では、ストレージ制御装置101は、HDD102に対してコマンドを送信したことに応じて、コマンドの実行結果を受信するとともに、メディアキャッシュ103の使用状況が変化したときに限り使用情報111を受信するものとする。

40

【0028】

図1の例では、ストレージ制御装置101は、HDD102#0から使用情報111#0を受信し、HDD102#1から使用情報111#1を受信し、...、HDD102#Nから使用情報111#Nを受信する。また、図1に示す使用情報111#0は、「Free」を示し、図1に示す使用情報111#1は、「Busy」を示すものとする。

【0029】

そして、HDD102へのアンアラインドとなる書き込みデータが発生したとする。この際、図1の(2)で示すように、ストレージ制御装置101は、使用情報111に基づいて、書き込みデータの書き込み要求をHDD102に送信する処理、または、自装置に

50

よる書き込みデータに対するR M W処理のいずれか一方を実行する。ここで、書き込みデータが発生する場合とは、例えば、ストレージ制御装置101が、ストレージ制御装置101の上位となるサーバから書き込み要求を受け付けた場合である。また、HDD#0~#Nのうち、いずれかのHDDのデータを、他のHDDに移動する場合にも、書き込みデータが発生することになる。

【0030】

図1の例では、ストレージ制御装置101の上位となるサーバから仮想的なディスクへの書き込み要求があったとする。そして、ストレージ制御装置101が書き込み先を判断した結果、HDD#0への書き込みデータwd#0__uとHDD#1への書き込みデータwd#1__uとが発生したとする。ここで、書き込みデータwd#0__u、wd#1__uは、ともに、アンアラインドであるとする。

10

【0031】

書き込みデータwd#0__uについて、使用情報111#0が「Free」であるので、ストレージ制御装置101は、書き込みデータwd#0__uの書き込み要求となるライトコマンドを、HDD102#0に送信する。ライトコマンドを受け付けたHDD102#0は、書き込みデータwd#0__uに対するR M W処理を実行し、アラインドとなる書き込みデータwd#0__aを得る。そして、HDD102#0は、書き込みデータwd#0__aを磁気ディスクに書き込む。

【0032】

また、書き込みデータwd#1__uについて、使用情報111#1が「Busy」であるので、ストレージ制御装置101は、書き込みデータwd#1__uに対するR M W処理を実行し、アラインドとなる書き込みデータwd#1__aを得る。そして、ストレージ制御装置101は、書き込みデータwd#1__aの書き込み要求となるライトコマンドをHDD102#1に送信する。ライトコマンドを受け付けたHDD102#1は、書き込みデータwd#1__aを磁気ディスクに書き込む。

20

【0033】

なお、図1における書き込みデータwd#0__a、wd#1__aについて、塗りつぶした領域が、それぞれ、書き込みデータwd#0__u、wd#1__uと同一のデータであるところを示し、白抜きの領域が、磁気ディスクから読み出したデータであることを示す。

【0034】

このように、ストレージ制御装置101は、2種類のR M W処理を最適化することにより、HDD102への書き込み時の性能を向上させることができる。また、ストレージ制御装置101は、HDD102に対して性能テストを行わないため、HDD102にデータを格納できない時間を発生させないようにすることができる。また、ストレージ制御装置101は、HDD102に対して性能テストを行わないため、性能テストにより発生したHDD102の寿命を縮めることを抑止することができる。次に、ストレージ制御装置101を、ディスクアレイ装置に適用した例を、図2を用いて説明する。

30

【0035】

図2は、情報処理システム200の一例を示す説明図である。情報処理システム200は、管理装置201と、ディスクアレイ装置202と、サーバ203とを含む。ディスクアレイ装置202およびサーバ203は、SAN(Storage Area Network)204に接続される。ディスクアレイ装置202、管理装置201およびサーバ203は、ネットワーク205に接続される。ネットワーク205は、例えばLAN(Local Area Network)である。

40

【0036】

ディスクアレイ装置202は、複数のHDDを内蔵し、サーバ203の処理に用いられるデータを記憶する。ディスクアレイ装置202は、データのライトコマンドをサーバ203から受信すると、内蔵のHDDにデータを書き込む。ディスクアレイ装置202は、書き込み結果をサーバ203に送信する。また、ディスクアレイ装置202は、データのリードコマンドをサーバ203から受信すると、内蔵のHDDからデータを読み出し、読

50

み出したデータをサーバ203に送信する。ディスクアレイ装置202は、HDDと併せてSSDを内蔵してもよい。

【0037】

管理装置201は、ディスクアレイ装置202の運用管理を行うサーバコンピュータである。管理装置201は、例えば、ネットワーク205を介してディスクアレイ装置202に動作制御の設定を行う。管理装置201は、ディスクアレイ装置202により実行されるプログラム、例えば、ファームウェアのプログラムを、ネットワーク205を介してディスクアレイ装置202に送信することがある。

【0038】

サーバ203は、ディスクアレイ装置202に格納されたデータにアクセスするサーバコンピュータである。サーバ203は、業務用のアプリケーションのプログラムを実行し、業務処理サービスをユーザに提供する。

【0039】

図3は、ディスクアレイ装置202のハードウェア構成例を示す説明図である。ディスクアレイ装置202は、コントローラモジュール(CM: Controller Module)301、302とDE(Drive Enclosure)303とを有する。

【0040】

CM301、302は、DE303が有するHDD102の記憶領域の管理や記憶領域に対するデータアクセスを制御する。CM301、302はディスクアレイ装置202内で冗長化され、データへのアクセス性能や耐故障性の向上が図られる。CM301、302は、図1に示したストレージ制御装置101の一例である。以下の例では、CM301について説明する。

【0041】

CM301は、プロセッサ311、RAM312、フラッシュメモリ313、CA(Channel Adaptor)314、NA(Network Adaptor)315およびDI(Drive Interface)316を有する。プロセッサ311~DI316は、バス317に接続される。CM302もCM301と同様のハードウェアを有する。

【0042】

プロセッサ311は、CM301の情報処理を制御するプロセッサである。プロセッサ311は、マルチプロセッサであってもよい。プロセッサ311は、例えばCPU(Central Processing Unit)、DSP(Digital Signal Processor)、ASIC(Application Specific Integrated Circuit)またはFPGA(Field Programmable Gate Array)などである。プロセッサ311は、CPU、DSP、ASIC、FPGAなどのうちの2以上の要素の組み合わせであってもよい。

【0043】

RAM312は、CM301の主記憶装置である。RAM312は、プロセッサ311に実行させるファームウェアのプログラムの少なくとも一部を一時的に記憶する。フラッシュメモリ313は、CM301の補助記憶装置である。フラッシュメモリ313は、不揮発性の半導体メモリである。フラッシュメモリ313は、ファームウェアのプログラムなどを記憶する。

【0044】

CA314は、SAN204を介してサーバ203と通信する通信インタフェースである。CA314の通信インタフェースの種別は、例えばファイバチャネルやSCSI(Small Computer System Interface)などである。CM301は、複数個のCAを有して、CAを冗長構成としてもよい。

【0045】

NA315は、ネットワーク205を介して他のコンピュータと通信する通信インタフェースである。NA315の通信インタフェースの種別は、例えばLAN接続用のNIC

10

20

30

40

50

(Network Interface Card)である。CM301は、複数個のNAを有して、NAを冗長構成としてもよい。

【0046】

DI316は、DE303と通信する通信インタフェースである。DI316の通信インタフェースの種別は、例えばファイバチャネルやSCSIなどである。CM301は、複数個のDIを有して、DIを冗長構成としてもよい。

【0047】

DE303は、HDD102#0～#Nを含む複数のHDDを収容する筐体である。Nは、1以上の整数である。そして、CM301、302は、HDD102#0～#Nを用いてRAID技術により、データへのアクセス性能や各HDDの耐障害性を向上した論理的な記憶領域を実現する。

10

【0048】

図4は、管理装置201のハードウェア構成例を示す説明図である。管理装置201は、プロセッサ401と、RAM402と、HDD403と、画像信号処理部404と、入力信号処理部405と、読み取り装置406と、通信インタフェース407とを有する。プロセッサ401～通信インタフェース407は管理装置201のバス408に接続される。

【0049】

プロセッサ401は、管理装置201の情報処理を制御する。プロセッサ401は、マルチプロセッサであってもよい。プロセッサ401は、例えばCPU、DSP、ASICまたはFPGAなどである。プロセッサ401は、CPU、DSP、ASIC、FPGAなどのうちの2以上の要素の組み合わせであってもよい。

20

【0050】

RAM402は、管理装置201の主記憶装置である。RAM402は、プロセッサ401に実行させるOS(Operating System)のプログラムやアプリケーションプログラムの少なくとも一部を一時的に記憶する。また、RAM402は、プロセッサ401による処理に用いる各種データを記憶する。

【0051】

HDD403は、管理装置201の補助記憶装置である。HDD403は、内蔵した磁気ディスクに対して、磁氣的にデータの書き込みおよび読み出しを行う。HDD403は、OSのプログラム、アプリケーションプログラム、および各種データを記憶する。管理装置201は、フラッシュメモリやSSDなどの他の種類の補助記憶装置を有してもよく、複数の補助記憶装置を有してもよい。

30

【0052】

画像信号処理部404は、プロセッサ401からの命令に従って、管理装置201に接続されたディスプレイ411に画像を出力する。ディスプレイ411としては、CRT(Cathode Ray Tube)ディスプレイや液晶ディスプレイなどを用いることができる。

【0053】

入力信号処理部405は、管理装置201に接続された入力デバイス412から入力信号を取得し、プロセッサ401に出力する。入力デバイス412としては、例えば、マウスやタッチパネルなどのポインティングデバイス、キーボードなどを用いることができる。

40

【0054】

読み取り装置406は、記録媒体413に記録されたプログラムやデータを読み取る装置である。記録媒体413として、例えば、フレキシブルディスク(FD:Flexible Disk)やHDDなどの磁気ディスク、CD(Compact Disc)やDVD(Digital Versatile Disc)などの光ディスク、光磁気ディスク(MO:Magneto-Optical disk)を使用することができる。また、記録媒体413として、例えば、フラッシュメモリカードなどの不揮発性の半導体メ

50

メモリを使用することもできる。読み取り装置406は、例えば、プロセッサ401からの命令に従って、記録媒体413から読み取ったプログラムやデータをRAM402またはHDD403に格納する。プロセッサ401は、記録媒体413から読み取ったプログラムやデータ、例えば、ディスクアレイ装置202のファームウェアのプログラムやデータをディスクアレイ装置202に送信することもできる。

【0055】

通信インタフェース407は、ネットワーク205を介して他の装置と通信を行う。通信インタフェース407は、有線通信インタフェースでもよいし、無線通信インタフェースでもよい。

【0056】

図5は、HDD102#0の例を示す説明図である。HDD102#0は、プロセッサ501と、メディアキャッシュ103と、磁気ディスク502__0~Mとを有する。Mは0以上の整数である。HDD102#1~#NもHDD102#0と同様のハードウェアを有する。

【0057】

プロセッサ501は、HDD102#0におけるRMWを実行する。プロセッサ501は、RMW用途のASICやFPGAなどでもよいし、RMW用のプログラムを実行する汎用プロセッサでもよい。

【0058】

メディアキャッシュ103は、プロセッサ501によるRMW処理に用いられる記憶領域である。メディアキャッシュは、RMW処理において、データを一時的に記憶するキャッシュとして利用される。メディアキャッシュ103は、HDD102#0が有する磁気ディスク502__0~Mの記憶領域の一部でもよいし、磁気ディスク502とは別個に設けられた半導体メモリでもよい。RMW用のプロセッサ501およびメディアキャッシュ103の実装方法は、HDD102のベンダによって異なるものとなる。例えば、メディアキャッシュ103が、磁気ディスク502__0~Mの記憶領域の一部であり、複数の領域に分割して存在してもよい。また、メディアキャッシュ103のデータサイズは、HDD102のベンダによって様々なものとなる。また、メディアキャッシュ103が存在しないHDD102も有り得る。

【0059】

図5では、磁気ディスク502__0を例として、磁気ディスク502__0の内部の例を示す。磁気ディスク502__1~Mも磁気ディスク502__0と同様である。磁気ディスク502__1上の同心円の縁に形成される領域はトラックと呼ばれる。図5で示すトラックTrは、磁気ディスク502__0上における複数のトラックのうちの1つである。トラックTr上を所定サイズ単位に物理的に区切った領域は、物理セクタと呼ばれる。物理セクタScは、トラックTr上における複数の物理セクタのうちの1つである。

【0060】

HDD102#1は、Advanced Formatに対応しているため、1つの物理セクタのサイズが、4096バイト=4KBとなる。ここで、磁気ディスク502上に4KBの物理セクタが配置された構造を4KBのアライメント(Alignment)と呼称することもある。また、1つの物理セクタのサイズは、8KB、16KBなど、他のサイズでもよい。この場合、HDD102#0の磁気ディスク502に対するデータ書き込みの操作は、4KBの物理セクタ単位に行われる。すなわち、当該物理セクタがHDD102#0におけるデータの書き込み単位となる。

【0061】

ここで、サーバ203で動作するOSなどのソフトウェアによっては、セクタサイズを512バイトとしてデータ書き込みを行うものもある。このようなソフトウェアからのライトコマンドでは、512バイトの整数倍の書き込みが要求される。すると、HDD102#0における物理セクタの4KBアライメントと整合しない書き込みが生じ得る。このような、4KBアライメントと整合しない書き込みが要求された場合に、プロセッサ50

10

20

30

40

50

1 は、R M W 処理を実行する。

【 0 0 6 2 】

(C M 3 0 1 の機能構成例)

図 6 は、C M 3 0 1 の機能構成例を示す説明図である。C M 3 0 1 は、制御部 6 0 0 を有する。制御部 6 0 0 は、応答結果受信部 6 0 1 と、判定部 6 0 2 と、R M W 実行部 6 0 3 と、送信部 6 0 4 と、を含む。制御部 6 0 0 は、記憶装置に記憶されたプログラムをプロセッサ 3 1 1 が実行することにより、各部の機能を実現する。記憶装置とは、具体的には、例えば、図 3 に示した R A M 3 1 2、フラッシュメモリ 3 1 3 などである。また、各部の処理結果は、プロセッサ 3 1 1 のレジスタや、プロセッサ 3 1 1 のキャッシュメモリ、R A M 3 1 2 等に格納される。また、C M 3 0 2 も、C M 3 0 1 と同等な機能を有する。

10

【 0 0 6 3 】

また、C M 3 0 1 は、メディアキャッシュ状態管理テーブル 6 1 1 にアクセス可能である。メディアキャッシュ状態管理テーブル 6 1 1 の情報は、R A M 3 1 2 といった記憶装置に格納されている。メディアキャッシュ状態管理テーブル 6 1 1 は、H D D 1 0 2 # 0 ~ # N の使用情報 1 1 1 を記憶するテーブルである。メディアキャッシュ状態管理テーブル 6 1 1 の設定内容の一例は、図 1 1 で示す。

【 0 0 6 4 】

応答結果受信部 6 0 1 は、H D D 1 0 2 # 0 ~ # N から使用情報 1 1 1 を受信する。また、受信した使用情報 1 1 1 は、メディアキャッシュ状態管理テーブル 6 1 1 に設定される。

20

【 0 0 6 5 】

判定部 6 0 2 は、H D D 1 0 2 # 0 ~ # N のいずれかへの書き込みデータが発生した際に、書き込みデータがアラインドかアンアラインドかを判定する。例えば、判定部 6 0 2 は、書き込みデータの先頭のアドレスと、書き込みデータのデータ長とが、ともに所定値で割り切れれば、書き込みデータがアンアラインドであると判定する。

【 0 0 6 6 】

制御部 6 0 0 は、判定部 6 0 2 が H D D 1 0 2 # 0 ~ # N のいずれかへのアンアラインドとなる書き込みデータが発生したと判定した場合に、使用情報 1 1 1 に基づいて、送信部 6 0 4 による処理、または、R M W 実行部 6 0 3 による処理のいずれか一方を実行する。

30

【 0 0 6 7 】

例えば、H D D 1 0 2 # 0 ~ # N のいずれかへのアンアラインドとなる書き込みデータが発生した際に、書き込み先の H D D 1 0 2 の使用情報 1 1 1 が「B u s y」であれば、R M W 実行部 6 0 3 は、書き込みデータに対する R M W 処理を実行する。R M W 処理によって得られたアラインドとなる書き込みデータは、書き込み先の H D D 1 0 2 に送信される。

【 0 0 6 8 】

また、H D D 1 0 2 # 0 ~ # N のいずれかへのアンアラインドとなる書き込みデータが発生した際に、書き込み先の H D D 1 0 2 の使用情報 1 1 1 が「F r e e」であれば、送信部 6 0 4 は、書き込みデータの書き込み要求を書き込み先の H D D 1 0 2 に送信する。これにより、書き込み先となる H D D 1 0 2 は、書き込みデータに対する R M W 処理を実行する。

40

【 0 0 6 9 】

次に、C M 3 0 1 による H D D 1 0 2 へのコマンド発行時の処理を示したフローチャートを、図 7 ~ 図 1 0 を用いて説明する。

【 0 0 7 0 】

図 7 は、ライトコマンド以外のコマンド発行時処理手順の一例を示すフローチャートである。ライトコマンド以外のコマンド発行時処理は、ライトコマンド以外のコマンドを C M 3 0 1 が発行する際に、C M 3 0 1 と H D D 1 0 2 とが行う処理である。

50

【 0 0 7 1 】

CM301が、コマンド対象のHDD102へライトコマンド以外のコマンドを発行する(ステップS701)。ここで、ライトコマンド以外のコマンドとは、例えば、リードコマンドや、テストユニットレディコマンド等がある。

【 0 0 7 2 】

コマンドを受け付けたコマンド対象のHDDが、CM301へコマンドの実行結果とメディアキャッシュ103の使用状況を示す識別子とを送信する(ステップS702)。実行結果と識別子とを受け付けたCMが、コマンド対象のHDDから受信した識別子を基にメディアキャッシュ状態管理テーブル611を更新する(ステップS703)。メディアキャッシュ状態管理テーブル611の更新例は、図11で示す。ステップS703の処理終了後、CM301とHDD102とは、ライトコマンド以外のコマンド発行時処理を終了する。

10

【 0 0 7 3 】

図8は、ライトコマンド発行時処理手順の一例を示すフローチャート(その1)である。また、図9は、ライトコマンド発行時処理手順の一例を示すフローチャート(その2)である。また、図10は、ライトコマンド発行時処理手順の一例を示すフローチャート(その3)である。ライトコマンド発行時処理は、CM301がライトコマンドをHDD102に発行する際に、CM301とHDD102とが行う処理である。なお、図9で示す各ステップは、ライトデータがアラインドである際の処理を示す。また、図10で示す各ステップは、ライトデータがアンアラインドであり、かつ、メディアキャッシュ103が枯渇していない場合の処理を示す。

20

【 0 0 7 4 】

CM301が、HDD102にこれから発行しようとしているライトコマンドのライトデータがアラインドかアンアラインドかを判定する(ステップS801)。そして、CM301が、ライトコマンドのライトデータがアンアラインドか否かを判断する(ステップS802)。ライトコマンドのライトデータがアンアラインドである場合(ステップS802:Yes)、CM301が、メディアキャッシュ状態管理テーブル611のうちのアクセス対象のHDD102のフィールドを参照する(ステップS803)。そして、参照結果に基づいて、CM301は、アクセス対象のHDD102のメディアキャッシュ103の状態がBusyか否かを判断する(ステップS804)。アクセス対象のHDD102のメディアキャッシュ103の状態がBusyである場合(ステップS804:Yes)、CM301が、ライトコマンドのライトデータに対するRMW処理を実行する(ステップS805)。

30

【 0 0 7 5 】

そして、アクセス対象のHDD102が、CM301へライトコマンドの実行結果とメディアキャッシュ103の使用状況を示す識別子とを送信する(ステップS806)。なお、ステップS806の処理において、CM301は、RMW処理により、リードコマンドとライトコマンドとをHDD102に送信することになる。従って、CM301は、RMW処理におけるリードコマンドの実行結果を受け付けた際に識別子を併せて受け付ける場合もあるし、RMW処理におけるライトコマンドの実行結果を受け付けた際に識別子を併せて受け付ける場合もある。

40

【 0 0 7 6 】

実行結果と識別子とを受け付けたCM301が、アクセス対象のHDD102から受信した識別子を基にメディアキャッシュ状態管理テーブル611を更新する(ステップS807)。ステップS807の処理終了後、CM301とHDD102とは、ライトコマンド発行時処理を終了する。

【 0 0 7 7 】

ライトコマンドのライトデータがアンアラインドでない場合(ステップS802:No)、CM301が、アラインドデータを含むライトコマンドをアクセス対象のHDD102に送信する(ステップS901)。ライトコマンドを受け付けたアクセス対象のHDD

50

102が、アラインドデータを磁気ディスク502に書き込む(ステップS902)。そして、アクセス対象のHDD102が、CM301へライトコマンドの実行結果とメディアキャッシュの使用状況を示す識別子とを送信する(ステップS903)。実行結果と識別子とを受け付けたCM301が、アクセス対象のHDD102から受信した識別子を基にメディアキャッシュ状態管理テーブル611を更新する(ステップS904)。ステップS904の処理終了後、CM301とHDD102とは、ライトコマンド発行時処理を終了する。

【0078】

アクセス対象のHDDのメディアキャッシュ103の状態がFreeである場合(ステップS804:No)、CM301が、アンアラインドデータを含むライトコマンドをアクセス対象のHDD102に送信する(ステップS1001)。ライトコマンドを受け付けたアクセス対象のHDD102が、RMW処理を実行する(ステップS1002)。そして、アクセス対象のHDDが、RMW処理によって得られたアラインドデータを磁気ディスク502に書き込む(ステップS1003)。次に、アクセス対象のHDDが、CM301へライトコマンドの実行結果とメディアキャッシュ103の使用状況を示す識別子とを送信する(ステップS1004)。

10

【0079】

実行結果と識別子とを受け付けたCM301が、アクセス対象のHDD102から受信した識別子を基にメディアキャッシュ状態管理テーブル611を更新する(ステップS1005)。ステップS1005の処理終了後、CM301とHDD102とは、ライトコマンド発行時処理を終了する。

20

【0080】

図11は、メディアキャッシュ状態管理テーブル611の記憶内容の一例を示す説明図である。図11では、メディアキャッシュ状態管理テーブル611の初期状態を示すメディアキャッシュ状態管理テーブル611aと、メディアキャッシュ状態管理テーブル611が初期状態から更新された状態を示すメディアキャッシュ状態管理テーブル611bとを示す。

【0081】

初期状態を示すメディアキャッシュ状態管理テーブル611aは、HDD102#0~#Nのそれぞれのメディアキャッシュ103がFreeであることを示す。そして、HDD102#1がアンアラインドデータのライトコマンドを受け付けたことによりBusyを応答した結果、メディアキャッシュ状態管理テーブル611が更新された状態が、メディアキャッシュ状態管理テーブル611bが示す状態である。メディアキャッシュ状態管理テーブル611bは、HDD102#1のメディアキャッシュ103がBusyであり、HDD102#0、#2~#Nのそれぞれのメディアキャッシュ103がFreeであることを示す。

30

【0082】

以上説明したように、CM301は、HDD102#0~#Nから受け付けた使用情報111#0~#Nに基づいて、RMW処理の主体をCM301と、書き込み先のHDD102とで切り替える。これにより、CM301は、メディアキャッシュ103の枯渇中にHDD102にRMW処理を行わせることを避けることができ、書き込み要求に対する応答性能の低下を抑えることができる。

40

【0083】

また、CM301は、CM301がコマンドをHDD102#0~#Nのいずれかに送信したことに応じて、コマンドを送信したHDD102からコマンドの実行結果を受け付けるとともに、使用情報111を受け付けてもよい。これにより、HDD102は、使用情報111の送信先を予め設定しなくてよい。例えば、図3で示したように、ディスクアレイ装置202内にはCMが複数ある。従って、HDD102が自発的に使用情報111を送信する場合には、ディスクアレイ装置202の管理者は、複数のCMのうちのいずれに使用情報111を送信するかを設定することになる。または、HDD102が、全ての

50

CMに使用情報111を送信するようにすると、ディスクアレイ装置202の負荷が増えることになる。

【0084】

また、HDD102は、CM301にコマンドの実行結果を送信する際にメディアキャッシュ103の使用状況が変化したときに限り使用情報111を送信しつつ、メディアキャッシュ103の使用状況が変化した際に使用情報111を自発的に送信してもよい。これにより、CM301が、メディアキャッシュ103の使用状況をより正確に把握することができて、書き込み要求に対する応答性能を向上することができる。

【0085】

例えば、CM301が、あるHDD102にライトコマンドを送信し、あるHDD102から、ライトコマンドの実行結果とともに、使用情報111として「Busy」を受け付けたとする。その後、あるHDD102は、RMW処理を完了させてメディアキャッシュ103が使用されていない状況となる。この状況となった後、あるHDD102は、使用情報111として「Free」をCM301に自発的に送信する。その後、CM301があるHDD102にいかなるコマンドも送信しないまま、あるHDD102に対してアンアラインドとなる書き込みデータが発生したとする。この場合、CM301は、メディアキャッシュ状態管理テーブル611内の使用情報111が「Free」であるため、あるHDD102にアンアラインドとなる書き込みデータを送信し、あるHDD102にRMW処理を行わせることができる。

【0086】

これに対し、RMW処理を完了させてメディアキャッシュ103が使用されていない状況となった後、あるHDD102が、使用情報111として「Free」をCM301に送信しなかったとする。そして、CM301があるHDD102にいかなるコマンドも送信しないまま、あるHDD102に対してアンアラインドとなる書き込みデータが発生したとする。この場合、CM301は、メディアキャッシュ状態管理テーブル611内の使用情報111が「Busy」であるため、CM301は、あるHDD102にRMW処理を行わせることが可能であるにも関わらず、自装置でRMW処理を行うことになる。

【0087】

また、CM301は、あるHDD102へのアンアラインドとなる書き込みデータが発生した際に、あるHDD102の使用情報111が「Busy」であれば、自装置によるRMW処理を実行してもよい。これにより、CM301は、メディアキャッシュ103の枯渇中にHDD102にRMW処理を行わせることを避けることができ、書き込み要求に対する応答性能の低下を抑えることができる。

【0088】

また、CM301は、あるHDD102へのアンアラインドとなる書き込みデータが発生した際に、あるHDD102の使用情報111が「Free」であれば、あるHDDにアンアラインドとなる書き込みデータの書き込み要求を送信してもよい。これにより、CM301は、HDD102にRMW処理を行わせることができ、CM301がRMW処理を実行する場合に比べて書き込み要求に対する応答性能を向上することができる。

【0089】

また、CM301は、あるHDD102への書き込みデータが発生した際に書き込みデータがアンアラインドであると判断した場合、使用情報111に基づいて、RMW処理の主体をCM301と、書き込み先のHDD102とで切り替えてもよい。このように、CM301は、書き込みデータがアンアラインドであるときに限り使用情報111を参照し、書き込みデータがアラインドであれば使用情報111を参照するといった無駄な処理を行わずにHDD102への書き込みを行うことができる。

【0090】

また、本実施の形態では、ストレージがHDDであることが好ましい。Advanced Formatは、HDDの大容量化および1度に処理されるデータ量の増大化に伴うものであるためである。また、物理セクタのサイズが8KBや16KB等に拡張されたF

10

20

30

40

50

o r m a t が出現した場合に、本実施の形態を適用することができる。また、ストレージが S S D であっても、S S D の物理セクタのサイズが 8 K B や 1 6 K B 等に拡張された F o r m a t が出現した場合等に、本実施の形態を適用することができる。

【 0 0 9 1 】

また、本実施の形態は、H D D # 0 ~ # N が、アクセス性能が異なるものであって、H D D # 0 ~ # N に対してストレージ階層化が適用されたシステムに対しても適用することができる。例えば、ストレージ階層化を管理する装置が、本実施の形態を適用して、ある階層を実現する H D D の状態情報に基づいて、H D D による R M W 処理、または、ストレージ階層化を管理する装置による R M W 処理を動的に切り替えてもよい。

【 0 0 9 2 】

なお、本実施の形態で説明したストレージ制御方法は、予め用意されたプログラムをパーソナル・コンピュータやワークステーション等のコンピュータで実行することにより実現することができる。本ストレージ制御プログラムは、ハードディスク、フレキシブルディスク、C D - R O M (C o m p a c t D i s c - R e a d O n l y M e m o r y)、D V D 等のコンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行される。また本ストレージ制御プログラムは、インターネット等のネットワークを介して配布してもよい。

【 0 0 9 3 】

上述した実施の形態に関し、さらに以下の付記を開示する。

【 0 0 9 4 】

(付記 1) アンアラインドとなる書き込みデータの書き込み要求に応じて前記書き込みデータに対するリードモディファイライト処理を行うストレージから受信した、前記リードモディファイライト処理に用いる記憶領域の使用状況を示す使用情報を記憶する記憶部と、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報に基づいて、当該書き込みデータの書き込み要求を前記ストレージに送信する処理、または、自装置による当該書き込みデータに対するリードモディファイライト処理のいずれか一方を実行する制御部と、

を有することを特徴とするストレージ制御装置。

【 0 0 9 5 】

(付記 2) 前記制御部は、

自装置がコマンドを前記ストレージに送信したことに応じた前記コマンドの実行結果とともに前記使用情報を受信する、

ことを特徴とする付記 1 に記載のストレージ制御装置。

【 0 0 9 6 】

(付記 3) 前記制御部は、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報によって前記記憶領域が使用されている状況が示される場合、自装置による当該書き込みデータに対するリードモディファイライト処理を実行する、

ことを特徴とする付記 1 または 2 に記載のストレージ制御装置。

【 0 0 9 7 】

(付記 4) 前記制御部は、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報によって前記記憶領域が使用されていない状況が示される場合、当該書き込みデータの書き込み要求を前記ストレージに送信する処理を実行する、

ことを特徴とする付記 1 ~ 3 のいずれか一つに記載のストレージ制御装置。

【 0 0 9 8 】

(付記 5) 前記制御部は、

前記ストレージへの書き込みデータが発生した際に、前記書き込みデータがアラインドかアンアラインドかを判定し、

10

20

30

40

50

前記書き込みデータがアンアラインドであると判定した場合、前記使用情報に基づいて、当該書き込みデータの書き込み要求を前記ストレージに送信する処理、または、自装置による当該書き込みデータに対するリードモディファイライト処理のいずれか一方を実行する、

ことを特徴とする付記 1 ~ 4 のいずれか一つに記載のストレージ制御装置。

【 0 0 9 9 】

(付記 6) 前記ストレージは、ハードディスクドライブである、

ことを特徴とする付記 1 ~ 5 のいずれか一つに記載のストレージ制御装置。

【 0 1 0 0 】

(付記 7) アンアラインドとなる書き込みデータの書き込み要求に応じて前記書き込みデータに対するリードモディファイライト処理を行うストレージと、前記ストレージに接続されたストレージ制御装置とを含むシステムであって、

前記ストレージは、

前記リードモディファイライト処理に用いる記憶領域の使用状況を示す使用情報を前記ストレージ制御装置に送信し、

前記ストレージ制御装置は、

前記ストレージから受信した、前記使用情報を記憶部に記憶させ、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報に基づいて、当該書き込みデータの書き込み要求を前記ストレージに送信する処理、または、自装置による当該書き込みデータに対するリードモディファイライト処理のいずれか一方を実行する、

ことを特徴とするシステム。

【 0 1 0 1 】

(付記 8) コンピュータが、

アンアラインドとなる書き込みデータの書き込み要求に応じて前記書き込みデータに対するリードモディファイライト処理を行うストレージから受信した、前記リードモディファイライト処理に用いる記憶領域の使用状況を示す使用情報を記憶部に記憶させ、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報に基づいて、当該書き込みデータの書き込み要求を前記ストレージに送信する処理、または、自装置による当該書き込みデータに対するリードモディファイライト処理のいずれか一方を実行する、

処理を実行することを特徴とするストレージ制御方法。

【 0 1 0 2 】

(付記 9) コンピュータに、

アンアラインドとなる書き込みデータの書き込み要求に応じて前記書き込みデータに対するリードモディファイライト処理を行うストレージから受信した、前記リードモディファイライト処理に用いる記憶領域の使用状況を示す使用情報を記憶部に記憶させ、

前記ストレージへのアンアラインドとなる書き込みデータが発生した際に、前記使用情報に基づいて、当該書き込みデータの書き込み要求を前記ストレージに送信する処理、または、自装置による当該書き込みデータに対するリードモディファイライト処理のいずれか一方を実行する、

処理を実行させることを特徴とするストレージ制御プログラム。

【符号の説明】

【 0 1 0 3 】

1 0 1 ストレージ制御装置

1 0 2 H D D

1 0 3 メディアキャッシュ

3 0 1、3 0 2 C M

6 0 0 制御部

6 0 1 応答結果受信部

10

20

30

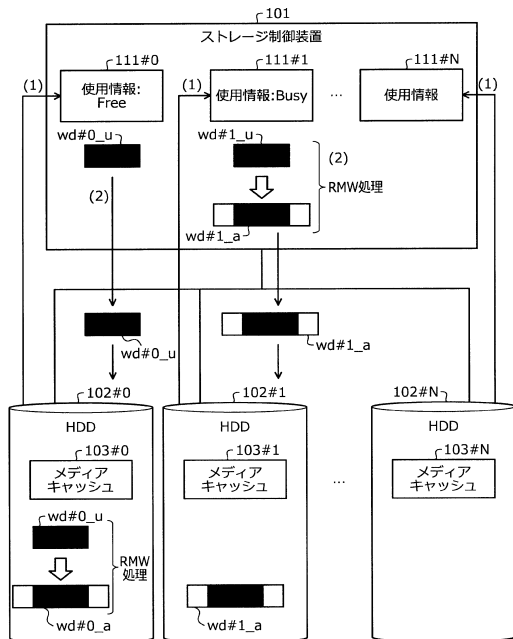
40

50

- 6 0 2 判定部
- 6 0 3 R M W 実行部
- 6 0 4 送信部
- 6 1 1 メディアキャッシュ状態管理テーブル

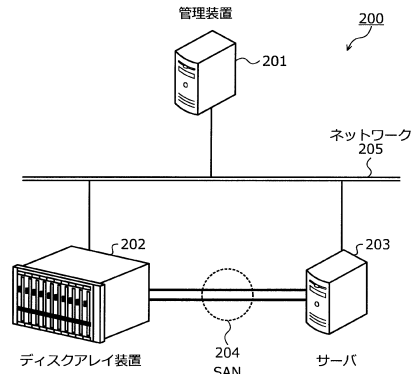
【 図 1 】

本実施の形態にかかるストレージ制御装置101の動作例を示す説明図



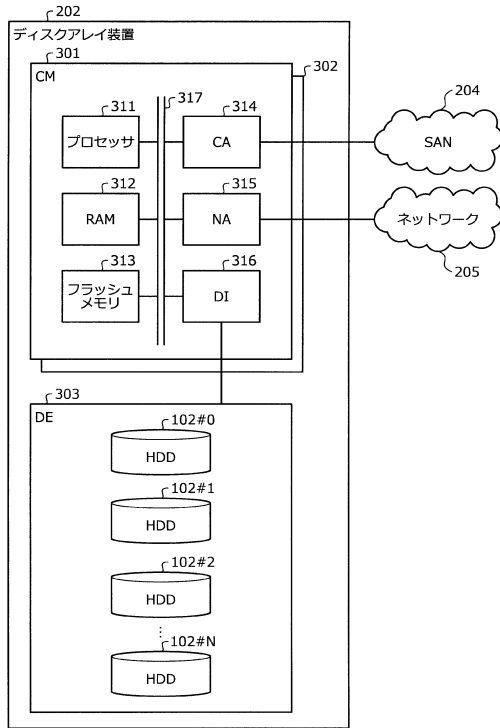
【 図 2 】

情報処理システム200の一例を示す説明図



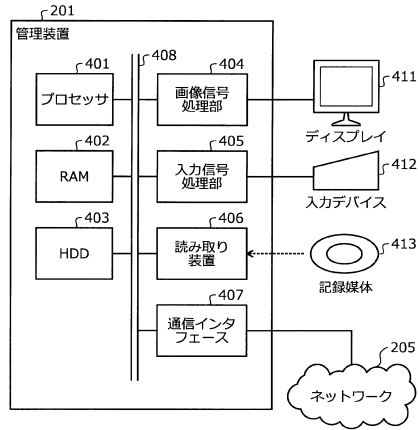
【図3】

ディスクレイ装置202のハードウェア構成例を示す説明図



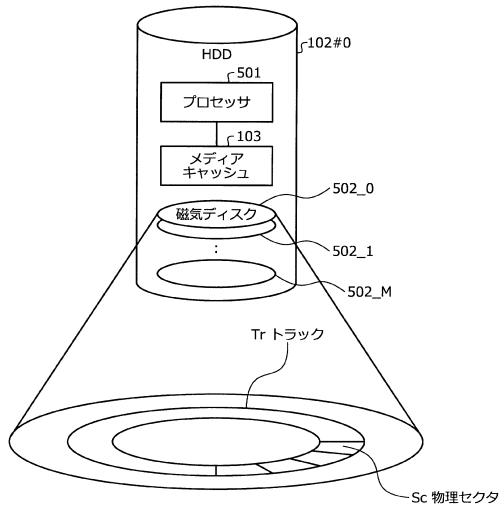
【図4】

管理装置201のハードウェア構成例を示す説明図



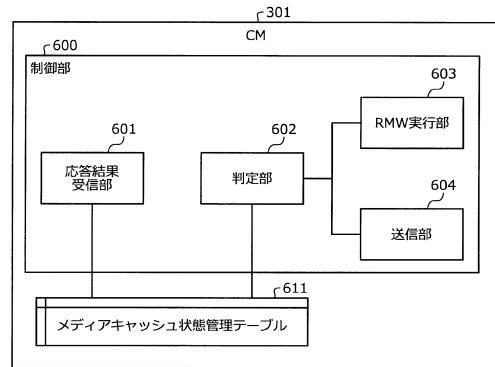
【図5】

HDD102#0の例を示す説明図



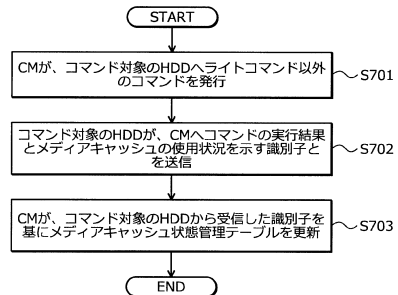
【図6】

CM301の機能構成例を示す説明図



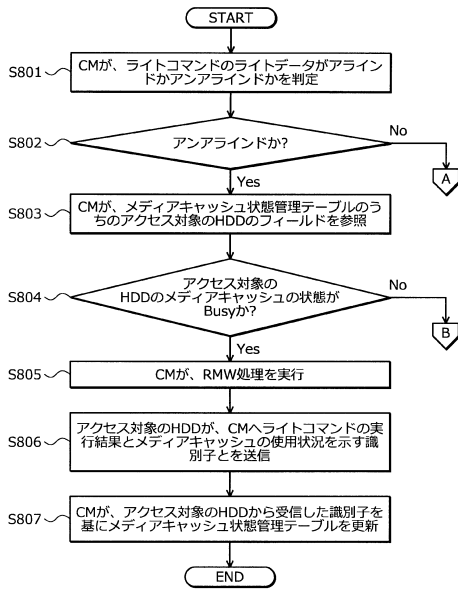
【図7】

ライトコマンド以外のコマンド発行時処理手順の一例を示すフローチャート



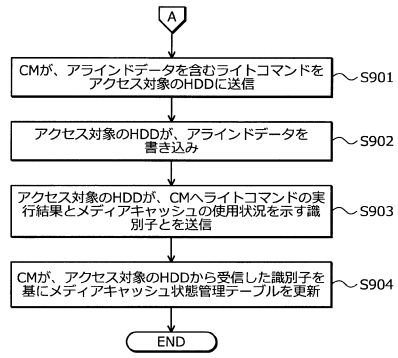
【 図 8 】

ライトコマンド発行時処理手順の一例を示すフローチャート(その1)



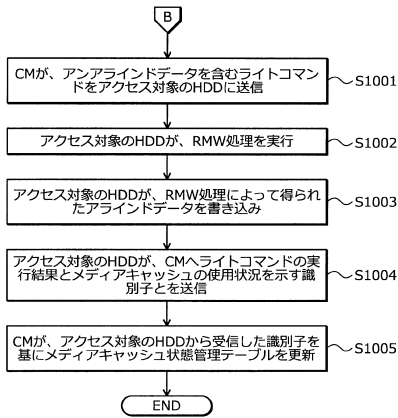
【 図 9 】

ライトコマンド発行時処理手順の一例を示すフローチャート(その2)



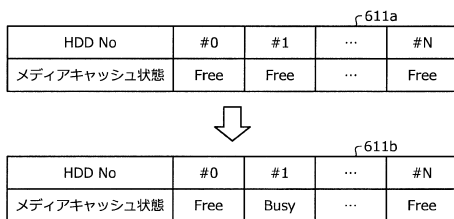
【 図 10 】

ライトコマンド発行時処理手順の一例を示すフローチャート(その3)



【 図 11 】

メディアキャッシュ状態管理テーブル611の記憶内容の一例を示す説明図



フロントページの続き

- (56)参考文献 特開2015-082269(JP,A)
特開2013-182373(JP,A)
特開2010-033396(JP,A)
特開2008-250919(JP,A)
特開2016-151850(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F3/06-3/08
G06F12/00-12/128
G06F13/10-13/14
G11B20/10-20/16