

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5696483号
(P5696483)

(45) 発行日 平成27年4月8日(2015.4.8)

(24) 登録日 平成27年2月20日(2015.2.20)

(51) Int.Cl.	F I
G 1 1 B 20/18 (2006.01)	G 1 1 B 20/18 5 1 2 D
G 0 6 F 3/06 (2006.01)	G 0 6 F 3/06 3 0 5 F
G 1 1 B 20/12 (2006.01)	G 0 6 F 3/06 3 0 6 B
G 1 1 B 20/10 (2006.01)	G 1 1 B 20/18 5 2 0 E
	G 1 1 B 20/18 5 5 2 B
	請求項の数 3 (全 28 頁) 最終頁に続く

(21) 出願番号	特願2011-3704 (P2011-3704)	(73) 特許権者	000005223 富士通株式会社
(22) 出願日	平成23年1月12日 (2011.1.12)		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2012-146359 (P2012-146359A)	(74) 代理人	100094330 弁理士 山田 正紀
(43) 公開日	平成24年8月2日 (2012.8.2)	(74) 代理人	100109689 弁理士 三上 結
審査請求日	平成25年10月7日 (2013.10.7)	(72) 発明者	齋藤 博紀 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
		審査官	堀 洋介
			最終頁に続く

(54) 【発明の名称】 情報記憶システム

(57) 【特許請求の範囲】

【請求項1】

情報の正誤を確認可能な符号を含んだ記憶情報を各々が記憶する複数の記憶区画が並んだトラックが、互いに隣接して複数本設けられた記憶媒体と、

前記記憶区画を単位としてアクセス範囲が指定され、前記記憶媒体上の該アクセス範囲にアクセスすることで前記記憶情報を読み出す読出部と、

前記読出部により読み出された前記記憶情報から正確な情報が得られたか否かを判定する判定部と、

正確な情報が得られなかったと前記判定部により判定された場合に前記読出部に、前記アクセス範囲からの前記記憶情報の読出しを繰り返させる繰返し指示部と、

正確な情報が得られたと前記判定部により判定された場合には、該正確な情報およびアクセス完了の通知の少なくとも一方を出力し、前記記憶情報の読出しが、予め決められた限度まで繰り返されても、正確な情報が得られなかったと前記判定部により判定された場合には、読出しの失敗を示すエラー情報を出力する第1出力部と、

特定の指示を受けた場合に、前記読出部により読み出された前記記憶情報の生データをそのまま出力する第2出力部と、

を備えた記憶装置；および

前記記憶装置に前記アクセス範囲を指定する指定部と、

前記記憶装置の前記第1出力部から出力された情報を取得し、その取得した情報が前記エラー情報である場合に、情報読出が試行された記憶区画である試行区画を含んだ第1ア

アクセス範囲と、該試行区画のトラックに隣接するトラック上の、該試行区画に隣接する記憶区画である隣接区画を含んだ第2アクセス範囲とを前記記憶装置に指定して、該第1アクセス範囲と該第2アクセス範囲から前記記憶情報の生データの読み出しを実行させるとともに、前記第2出力部に対して前記特定の指示を与える指示付与部と、

前記第2出力部によりそのまま出力された前記記憶情報の生データを取得し、その取得した前記記憶情報の生データから正確な情報が、前記試行区画および前記隣接区画の双方について得られない場合、該試行区画に替えて前記読出部によるアクセス対象となる交代区画を割り当てる交代部と、

を備えた制御装置；

を備えたことを特徴とする情報記憶システム。

10

【請求項2】

前記指示付与部が、前記第2アクセス範囲として、該試行区画のトラックにそれぞれが隣接して相互間に該トラックを挟んだ2つのトラックそれぞれ上の隣接区画を含んだ2つの第2アクセス範囲を前記記憶装置に指定するものであり、

前記交代部が、前記試行区画と、該試行区画を挟んだ両方の前記隣接区画とのいずれについても正確な情報が得られない場合の該試行区画について前記交代区画を割り当てるものであることを特徴とする請求項1記載の情報記憶システム。

【請求項3】

前記交代部が、前記試行区画および前記隣接区画の双方について正確な情報が得られない場合の該試行区画および該隣接区画について前記交代区画を割り当てるものであることを特徴とする請求項2記載の情報記憶システム。

20

【発明の詳細な説明】

【技術分野】

【0001】

本件は、情報を記憶する情報記憶システムに関する。

【背景技術】

【0002】

近年、HDD(Hard Disk Drive)と、そのHDDに指示して情報の読書きを行わせる制御装置とを備えた情報記憶システムが普及している。

【0003】

30

HDDには、内蔵の磁気ディスクにおける例えば1個のセクタ等の記憶区画を単位としてアクセス範囲が指定される。そして、その指定されたアクセス範囲から情報が読み出される。また、多くの場合、各記憶区画が記憶している記憶情報は、その情報の正誤を確認可能ないわゆるECC(Error Checking and Correction)符号を含んだものとなっている。このECC符号は、各記憶区画から記憶情報が読み出されたときの、その読み出された記憶情報についての誤り検出と、その記憶情報から誤りが検出された場合の情報修正に用いられる。記憶区画から記憶情報が読み出されると、その記憶情報についての誤り検出が、その記憶情報に含まれているECC符号を使って行われる。そして、読み出された記憶情報から誤りが検出された場合には、その読み出された記憶情報の修正がECC符号を使って行われる。

40

【0004】

HDDでは、指定されたアクセス範囲の各記憶区画からの記憶情報の読出しは、そのアクセス範囲の先頭に位置する記憶区画から順次に行われる。

【0005】

また、ある記憶区画から読み出された記憶情報から誤りが検出されなかった場合、あるいは、誤りが検出されてもECC符号を使って修正された場合には、HDDでは、その記憶区画から読み出された記憶情報から正確な情報が得られたと判断される。ある記憶区画について、このように正確な情報が得られたと判断されると、次の記憶区画についての情報の読出し、誤り検出等が行われる。

【0006】

50

一方、ある記憶区画の記憶情報について、誤りが検出され、且つ、修正も出来なかった場合には、その記憶区画から記憶情報が再び読出されて、その再び読み出された記憶情報について誤り検出と修正がECC符号を使って行われる。この再度の読出し、誤り検出、修正という一連の処理を、以下、リトライ処理と呼ぶ。リトライ処理は、読み出された記憶情報が修正されることで、その記憶情報から正確な情報が得られたと判断されるまで繰り返される。そして、このようなリトライ処理を経て正確な情報が得られた場合も、次の記憶区画についての情報の読出し、誤り検出等が行われる。

【0007】

ここで、HDDの多くは、上記のリトライ処理に要した時間の合計が、予め決められた制限時間に達した場合には、情報の読出しを打ち切り、HDDの制御装置に向けて、HDD内部での読出しの失敗を示すエラー情報を出力する。

10

【0008】

上記の記憶区画から読み出された記憶情報から検出される誤りには、例えばヘッドで情報を読み出す際のクロストークやノイズ等に起因した軽微な誤りや、傷等の欠陥に起因した重篤な誤りがある。

【0009】

軽微な誤りが検出された場合には、少ない回数のリトライ処理で、正確な情報を得ることが可能であることが多い。一方で、重篤な誤りが検出された場合には、リトライ処理が何度も繰り返されても、記憶情報から正確な情報が得られないことが多い。

【0010】

20

そこで、HDDに、上記のエラー情報と共に、制限時間に達した時点でリトライ処理中であつた記憶区画を、欠陥が存在している記憶区画として制御装置に報告させ、その制御装置で次のような交代処理を行うことが知られている。

【0011】

交代処理は、HDDから報告されてきた記憶区画について、その記憶区画に替えて、アクセス対象として予備の記憶区画(交代区画)を割り当てる処理である(例えば、特許文献1および特許文献2参照)。また、このときには、交代区画には、元々の記憶区画で記憶しておくべき情報が書き込まれる。

【0012】

この交代処理によれば、制御装置に報告した記憶区画を含むアクセス範囲が指定されたときには、HDDでは、情報が、その記憶区画とは別の交代区画から読み出される。その結果、その報告された記憶区画についてのリトライ処理の発生が抑えられて、延いては、情報記憶システムとしてのパフォーマンスの低下が抑えられる。

30

【先行技術文献】

【特許文献】

【0013】

【特許文献1】特開平11-95933号公報

【特許文献2】特開平11-338648号公報

【発明の概要】

【発明が解決しようとする課題】

40

【0014】

ここで、HDDからエラー情報が出力されたときには、アクセス範囲内に、上記のように報告された記憶区画とは別に、何度もリトライ処理が繰り返された末に正確な情報が得られた記憶区画がある可能性がある。このように何度もリトライ処理が繰り返される記憶区画にも欠陥が存在している可能性が高い。しかし、エラー情報と共にHDDの制御装置に報告された記憶区画とは別にこのような記憶区画があつたとしても、制御装置では、その記憶区画の把握が困難であるのが現状である。そして、このような記憶区画が存在していると、その記憶区画を含むアクセス範囲が指定された場合に、また、その記憶区画についてリトライ処理が何度も繰り返される恐れがある。つまり、このような記憶区画の存在は、HDDとホスト装置とを有する情報記憶システムのパフォーマンス低下に繋がる恐れ

50

がある。

【 0 0 1 5 】

ここで、HDDにおける情報の読出しは、HDDの制御装置から情報の読出しが指示された場合だけでなく、情報の書込みが指示された場合にも、次のように行われる。

【 0 0 1 6 】

即ち、HDDでは、アクセス範囲が指定されて、そのアクセス範囲の記憶区画に情報が書き込まれた後に、その書込みが正常に完了したか否かの確認のために、そのアクセス範囲の各記憶区画に書き込まれてその記憶区画が記憶している記憶情報が読み出される。そして、その読出しに当たって、上述した誤り検出、修正、リトライ処理が行われる。

【 0 0 1 7 】

つまり、上記のようなリトライ処理の繰り返しによるパフォーマンス低下は、HDDの制御装置からHDDに読出しが指示された場合だけでなく、書込みが指示された場合にも起こり得る。

【 0 0 1 8 】

尚、ここまで、HDDを記憶装置として有する情報記憶システムを例に挙げて、リトライ処理の繰り返しによるパフォーマンス低下について説明した。

【 0 0 1 9 】

しかし、このようなパフォーマンス低下は、例えば光磁気ディスクやDVD(Digital Versatile Disc)等のドライブ装置を記憶装置として有する情報記憶システムについても同様に起こる得る。

【 0 0 2 0 】

本件は上記事情に鑑み、パフォーマンス低下が抑えられた情報記憶システムを提供することを目的とするものである。

【課題を解決するための手段】

【 0 0 2 1 】

上記目的を達成する情報記憶システムは、記憶装置と制御装置と備えている。

【 0 0 2 2 】

記憶装置は、記憶媒体と、読出部と、判定部と、繰り返し指示部と、第1出力部と、第2出力部とを備えている。

【 0 0 2 3 】

記憶媒体は、情報の正誤を確認可能な符号を含んだ記憶情報を各々が記憶する複数の記憶区画が並んだトラックが、互いに隣接して複数本設けられたものである。

【 0 0 2 4 】

読出部は、上記記憶区画を単位としてアクセス範囲が指定され、上記記憶媒体上のそのアクセス範囲にアクセスすることで上記記憶情報を読み出すものである。

【 0 0 2 5 】

判定部は、上記読出部により読み出された上記記憶情報から正確な情報が得られたか否かを判定するものである。

【 0 0 2 6 】

繰り返し指示部は、正確な情報が得られなかったと上記判定部により判定された場合に上記読出部に、上記アクセス範囲からの上記記憶情報の読出しを繰り返させるものである。

【 0 0 2 7 】

第1出力部は、正確な情報が得られたと上記判定部により判定された場合には、その正確な情報およびアクセス完了の通知の少なくとも一方を出力する。

【 0 0 2 8 】

一方、この第1出力部は、上記記憶情報の読出しが、予め決められた限度まで繰り返されても、正確な情報が得られなかったと上記判定部により判定された場合には、読出しの失敗を示すエラー情報を出力する。

【 0 0 2 9 】

第2出力部は、特定の指示を受けた場合に、上記読出部により読み出された上記記憶情

10

20

30

40

50

報をそのまま出力するものである。

【0030】

制御装置は、指定部と、指示付与部と、交代部とを備えている。

【0031】

指定部は、上記記憶装置に上記アクセス範囲を指定するものである。

【0032】

指示付与部は、上記記憶装置の上記第1出力部から出力された情報を取得し、その取得した情報が上記エラー情報である場合に、次のような処理を行う。

【0033】

この場合、指示付与部は、情報読出が試行された記憶区画である試行区画を含んだ第1アクセス範囲と、その試行区画のトラックに隣接するトラック上の、その試行区画に隣接する記憶区画である隣接区画を含んだ第2アクセス範囲とを上記記憶装置に指定する。そして、指示付与部は、上記第2出力部に対して上記特定の指示を与える。

10

【0034】

交代部は、上記第2出力部により出力された上記記憶情報を取得する。そして、交代部は、その取得した記憶情報から正確な情報が、上記試行区画および上記隣接区画の双方について得られない場合、次のような処理を行う。即ち、交代部は、上記試行区画に替えて上記読出部によるアクセス対象となる交代区画を割り当てる。

【発明の効果】

【0035】

本件によれば、パフォーマンスの低下が抑えられた情報記憶システムを得ることができる。

20

【図面の簡単な説明】

【0036】

【図1】情報記憶システムの第1実施形態を示す図である。

【図2】情報記憶システムの第2実施形態を示す図である。

【図3】図2に示すHDDの構成を示す図である。

【図4】図2に示すRAIDシステムに搭載されているHDDを、そのHDDの機能に注目して示す機能ブロック図である。

【図5】図2に示すRAIDシステムに搭載されているRAIDコントローラを、そのRAIDコントローラの機能に注目して示す機能ブロック図である。

30

【図6】RAIDコントローラで実行される処理の流れを示すフローチャートである。

【図7】HDDがエラー情報を出力する状況を模式的に示す図である。

【図8】アクセス範囲の各LBAについての、記憶情報が正当な情報であるか否かの判定結果の一例を模式的に示す図である。

【図9】マッピング処理を模式的に示す図である。

【発明を実施するための形態】

【0037】

以下、本件の情報記憶システムの具体的な実施形態について、図面を参照して説明する。

40

【0038】

まず、第1実施形態について説明する。

【0039】

図1は、情報記憶システムの第1実施形態を示す図である。

【0040】

図1に示す情報記憶システム1は、記憶装置10と制御装置20とを備えている。

【0041】

記憶装置10は、記憶媒体11と、読出部12と、判定部13と、繰返し指示部14と、第1出力部15と、第2出力部16とを備えている。

【0042】

50

記憶媒体 1 1 は、情報の正誤を確認可能な符号を含んだ記憶情報を各々が記憶する複数の記憶区画 1 1 a __ 1 が並んだトラック 1 1 a が、互いに隣接して複数本設けられたものである。

【 0 0 4 3 】

読出部 1 2 は、上記記憶区画 1 1 a __ 1 を単位としてアクセス範囲が指定され、上記記憶媒体 1 1 上のそのアクセス範囲にアクセスすることで上記記憶情報を読み出す。

【 0 0 4 4 】

尚、この読出部 1 2 が行う記憶情報の読出しには、この記憶装置 1 0 が読出し指示を受けた場合の読出しだけでなく、書込み指示を受けた場合の次のような読出しも含まれる。即ち、読出部 1 2 が行う読出しには、書込み指示を受けてアクセス範囲の記憶区画に情報が書き込まれた後に、その書込みが正常に完了したか否かの確認のために行われる記憶情報の読出しも含まれる。

10

【 0 0 4 5 】

判定部 1 3 は、上記読出部 1 2 により読み出された上記記憶情報から正確な情報が得られたか否かを判定する。

【 0 0 4 6 】

繰り返し指示部 1 4 は、正確な情報が得られなかったと上記判定部 1 3 により判定された場合に上記読出部 1 2 に、上記アクセス範囲からの上記記憶情報の読出しを繰り返させる。

【 0 0 4 7 】

20

第 1 出力部 1 5 は、正確な情報が得られたと上記判定部 1 3 により判定された場合には、その正確な情報およびアクセス完了の通知の少なくとも一方を出力する。

【 0 0 4 8 】

ここで、この第 1 出力部 1 5 は、この記憶装置 1 0 が読出し指示を受けていた場合には、上記の正確な情報およびアクセス完了の通知の両方を出力する。また、この記憶装置 1 0 が書込み指示を受けていた場合には、第 1 出力部 1 5 は、アクセス完了の通知のみを出力する。

【 0 0 4 9 】

一方、第 1 出力部 1 5 は、上記記憶情報の読出しが、予め決められた限度まで繰り返されても、正確な情報が得られなかったと上記判定部 1 3 により判定された場合には、読出しの失敗を示すエラー情報を出力する。

30

【 0 0 5 0 】

第 2 出力部 1 6 は、特定の指示を受けた場合に、上記読出部 1 2 により読み出された上記記憶情報をそのまま出力する。

【 0 0 5 1 】

制御装置 2 0 は、指定部 2 1 と、指示付与部 2 2 と、交代部 2 3 とを備えている。

【 0 0 5 2 】

指定部 2 1 は、上記記憶装置 1 0 に上記アクセス範囲を指定する。

【 0 0 5 3 】

指示付与部 2 2 は、上記記憶装置 1 0 の上記第 1 出力部 1 5 から出力された情報を取得し、その取得した情報が上記エラー情報である場合に、次のような処理を行う。

40

【 0 0 5 4 】

この場合、指示付与部 2 2 は、第 1 アクセス範囲と第 2 アクセス範囲とを上記記憶装置 1 0 に指定する。第 1 アクセス範囲は、情報読出が試行された記憶区画 1 1 a __ 1 である試行区画を含んだアクセス範囲である。第 2 アクセス範囲は、その試行区画のトラック 1 1 a に隣接するトラック 1 1 a 上の、その試行区画に隣接する記憶区画 1 1 a __ 1 である隣接区画を含んだアクセス範囲である。そして、指示付与部 2 2 は、上記第 2 出力部 1 6 に対して上記特定の指示を与える。

【 0 0 5 5 】

交代部 2 3 は、上記第 2 出力部 1 6 により出力された上記記憶情報を取得する。そして

50

、交代部 2 3 は、その取得した記憶情報から正確な情報が、上記試行区画および上記隣接区画の双方について得られない場合のその試行区画については、次のような処理を行う。即ち、交代部 2 3 は、上記試行区画については、上記記憶媒体 1 1 上の記憶区画 1 1 a _ 1 のうち、その試行区画に替えて上記読出部 1 2 によるアクセス対象となる交代区画を割り当てる。

【 0 0 5 6 】

例えば HDD において、ある記憶区画について上記のリトライ処理が何度も繰り返されて、やっと正確な情報が得られる場合には、その記憶区画に傷等の欠陥が存在している可能性が高い。そして、傷等の欠陥は、その記憶区画が属するトラックに隣接する隣のトラックの記憶区画にまで及んでいる可能性が高い。

10

【 0 0 5 7 】

図 1 に示す情報記憶システム 1 では、記憶装置 1 0 からエラー情報が出力されると、情報読出しが試行された上記試行区画を含んだ第 1 アクセス範囲の記憶区画と、上記隣接区画を含んだ第 2 アクセス範囲の記憶区画から情報がそのまま読み出される。それぞれのアクセス範囲の記憶区画からそのまま読み出された情報の中には、誤りを有する情報が含まれている。ここで、この誤りの中には、リトライ処理を何度も繰り返させる欠陥による重篤な誤りや、クロストークやノイズによる、少ない回数のリトライ処理で修正可能な軽微な誤り等といった、軽重様々なものがある。そして、上記のように、重篤な誤りに繋がる欠陥は、隣のトラックの記憶区画にまで及んでいる可能性が高い。

【 0 0 5 8 】

20

この情報記憶システム 1 では、上記試行区画と上記隣接区画との双方について正確な情報が得られない場合に、その試行区画については、交代区画が割り当てられる。これにより、互いに隣り合った複数のトラック間に亘る重大な誤りに繋がる欠陥が存在している可能性が高い記憶区画について交代区画が割り当てられることとなる。そして、次回に、この記憶区画を含むアクセス範囲からの情報の読出しが情報記憶システム 1 において実行されるときには、その記憶区画については交代区画にアクセスされることとなる。その結果、上記のようなリトライ処理の繰返し発生が抑えられる。つまり、この情報記憶システム 1 では、パフォーマンスの低下が抑えられることとなる。

【 0 0 5 9 】

次に、第 2 実施形態について説明する。

30

【 0 0 6 0 】

図 2 は、情報記憶システムの第 2 実施形態を示す図である。

【 0 0 6 1 】

この図 2 には、複数の HDD 1 0 0 を備え、情報を、これら複数の HDD 1 0 0 に記憶する RAID (Redundant Arrays of Inexpensive Disks) 構造を有する RAID システム 2 が示されている。

【 0 0 6 2 】

ここで、RAID 構造には、ある HDD に記憶した情報と同じ情報を別の HDD に記憶しておく RAID 1、情報を、その情報の修正のための情報とともに複数の HDD に分散して記憶する RAID 2 ~ 5 等といった複数種類がある。本実施形態の RAID システム 2 は、いずれの種類 RAID 構造であっても良いが、説明を簡単なものとするため、以下では、RAID 構造が RAID 1 であるものとして説明を行う。

40

【 0 0 6 3 】

図 2 示すように、RAID システム 2 は、複数の HDD 1 0 0 が収納されたドライブエンクロージャ 5 0 と、そのドライブエンクロージャ 5 0 内の HDD 1 0 0 に対して情報の読書きを指示する RAID コントローラ 7 0 とを備えている。

【 0 0 6 4 】

RAID コントローラ 7 0 は、HDD 1 0 0 に対する情報の読書き指示を含む各種処理を実行する制御 IC (Integrated Circuit) 7 1 を備えている。また、RAID コントローラ 7 0 は、ROM (Read Only Memory) 7 2 と、

50

R A M (R a n d o m A c c e s s M e m o r y) 7 3 を備えている。R O M 7 2 には、制御 I C 7 1 が実行する各種処理を記述した各種プログラムが格納されている。制御 I C 7 1 は、この R O M 7 2 内のプログラムに従って処理を実行する。また、R A M 7 3 には、R O M 7 2 中のプログラムが展開されると共に、処理の実行で使われる各種パラメータが格納される。さらに、R A I D コントローラ 7 0 は、書換え可能な不揮発性メモリとして半導体メモリ 7 4 を備えている。この半導体メモリ 7 4 には、後述の交代対応表が格納されている。

【 0 0 6 5 】

図 3 は、図 2 に示す H D D の構成を示す図である。

【 0 0 6 6 】

この図 3 に示すように、H D D 1 0 0 は、磁気ディスク 1 1 0、ヘッド 1 0 1、制御 I C 1 0 2、E C C 処理回路 1 0 3、R O M 1 0 4、R A M 1 0 5、および半導体メモリ 1 0 6 を備えている。

【 0 0 6 7 】

ここで、H D D 1 0 0 は、複数枚の磁気ディスク 1 1 0 を備えているが、枚数については特定せずに説明を行う。

【 0 0 6 8 】

磁気ディスク 1 1 0 には、各々が情報を記憶する複数のセクタ 1 1 1 a が周回して並んだトラック 1 1 1 が、互いに隣接して複数本設けられている。複数本のトラック 1 1 1 のそれぞれには、磁気ディスク 1 1 0 上での各トラック 1 1 1 の位置を示す番号が割り当てられている。また、各トラック 1 1 1 に属する複数のセクタ 1 1 1 a それぞれには、トラック 1 1 1 a 上での位置を示す番号が割り当てられている。

【 0 0 6 9 】

本実施形態における磁気ディスク 1 1 0 が、複数の記憶区画が並んだトラックが、互いに隣接して複数本設けられた記憶媒体の一例に相当する。また、この磁気ディスク 1 1 0 におけるセクタ 1 1 1 a が、上記の記憶媒体における記憶区画の一例に相当する。そして、この磁気ディスク 1 1 0 におけるトラック 1 1 1 a が、上記の記憶媒体におけるトラックの一例に相当する。

【 0 0 7 0 】

また、本実施形態では、磁気ディスク 1 1 0 は表裏両面に情報を記憶する。そして、H D D 1 0 0 は、各磁気ディスク 1 1 0 の表裏各面につき 1 つずつ、アクセスを行うヘッド 1 0 1 を備えている。そして、各ヘッド 1 0 1 には、そのヘッド 1 0 1 を示す番号が割り当てられている。

【 0 0 7 1 】

各 H D D 1 0 0 の内部では、磁気ディスク 1 1 0 上の各セクタ 1 1 1 a の位置が C H S (C y l i n d e r H e a d S e c t o r) 形式で取り扱われる。

【 0 0 7 2 】

C H S 形式は、いわゆる円筒座標系の一種で、各セクタ 1 1 1 a の位置を、シリンダ番号とヘッド番号とセクタ番号とで示す形式である。シリンダ番号は、同径トラックが並んだ円筒 (シリンダ) の番号である。ヘッド番号は、そのトラック 1 1 1 にアクセスするヘッド 1 0 1 の番号である。セクタ番号は、そのトラック 1 1 1 上でのセクタ 1 1 1 a の番号である。

【 0 0 7 3 】

一方で、図 2 の R A I D システム 2 では、各セクタ 1 1 1 a に、複数の H D D 1 0 0 に亘る通し番号である論理アドレス (L B A : L o g i c a l B l o c k A d d r e s s) が割り当てられている。以下、上記の C H S 形式で表されるセクタ 1 1 1 a の位置を物理アドレスと呼ぶ。R A I D システム 2 に接続される上位装置であるサーバ等には、各セクタ 1 1 1 a の位置を示すアドレスとして、この L B A が公開されている。

【 0 0 7 4 】

ここで、図 2 の R A I D システム 2 では、各 H D D 1 0 0 において、セクタ 1 1 1 a に

10

20

30

40

50

L B A が割り当てられていない予備領域が存在している。L B A は、各 H D D 1 0 0 の予備領域内のセクタ 1 1 1 a を除いたセクタ 1 1 1 a に割り当てられた通し番号となっている。一方で、物理アドレスは、各 H D D 1 0 0 における全てのセクタ 1 1 1 a に割り当てられている。

【 0 0 7 5 】

本実施形態では、上位装置から R A I D システム 2 への情報の読書き指示は、情報が読書きされる一連のセクタ 1 1 1 a を示すアクセス範囲が、上位装置から次のように L B A によって指定されて行われる。即ち、アクセス範囲として、上位装置からは、一連のセクタ 1 1 1 a のうち、アクセス順での先頭に位置するセクタ 1 1 1 a の L B A (先頭 L B A) と、その一連のセクタ 1 1 1 a の数 (ブロック数) とが指定される。また、情報の書込み指示の場合には、書込み情報が、この上位装置から R A I D システム 2 に送られる。

10

【 0 0 7 6 】

R A I D コントローラ 7 0 は、この上位装置からの読書き指示を受けて、R A I D システム 2 中の複数の H D D 1 0 0 のうち、上位装置から指定されたアクセス範囲のセクタ 1 1 1 a を有する H D D 1 0 0 に読書きを指示する。この指示の際には、R A I D コントローラ 7 0 は、その H D D 1 0 0 に、上位装置から指定された先頭 L B A とブロック数とでアクセス範囲を指定する。また、書込み指示の場合には、R A I D コントローラ 7 0 は、アクセス範囲の指定に加えて、上位装置から送られて来る書込み情報もその H D D 1 0 0 に送る。

【 0 0 7 7 】

20

制御 I C 1 0 2 は、R A I D コントローラ 7 0 から指定された先頭 L B A を物理アドレスに変換する。本実施形態では、R O M 1 0 4 に、H D D 1 0 0 内のセクタ 1 1 1 a についての、L B A から物理アドレスへの変換関係を表すアドレス変換情報が格納されている。制御 I C 1 0 2 による先頭 L B A の物理アドレスへの変換は、このアドレス変換情報を使って行われる。

【 0 0 7 8 】

ここで、上述したように、L B A は、H D D 1 0 0 内の、予備領域内のセクタ 1 1 1 a を除いたセクタ 1 1 1 a に割り当てられている。一方で、物理アドレスは、H D D 1 0 0 内の全てのセクタ 1 1 1 a に割り当てられている。上記のアドレス変換情報は、L B A から、予備領域内のセクタ 1 1 1 a を除いたセクタ 1 1 1 a への変換関係を表している。

30

【 0 0 7 9 】

尚、図 2 に示す R A I D コントローラ 7 0 の R O M 7 2 には、この R A I D システム 2 に搭載されている複数の H D D 1 0 0 それぞれについて、アドレス変換情報が格納されている。また、本実施形態では、複数の H D D 1 0 0 それぞれに識別番号が割り当てられている。R O M 7 2 内の各 H D D 1 0 0 のアドレス変換情報には、その H D D 1 0 0 の識別番号が付されている。

【 0 0 8 0 】

H D D 1 0 0 の制御 I C 1 0 2 は、先頭 L B A を、その H D D 1 0 0 の R O M 1 0 4 内のアドレス変換情報を使って、物理アドレスに変換する。また、情報の書込み時には、制御 I C 1 0 2 は、R A I D コントローラ 7 0 から送られて来る書込み情報をヘッド 1 0 1 に送る。そして、制御 I C 1 0 2 は、ヘッド 1 0 1 に、先頭 L B A を変換した物理アドレスのセクタ 1 1 0 a から、指定されたブロック数のセクタ 1 1 1 a に対して順次に情報を読書きさせる。

40

【 0 0 8 1 】

ここで、本実施形態では、各セクタ 1 1 1 a に書き込まれて、そのセクタ 1 1 1 a で記憶される記憶情報には、情報の正誤を確認可能ないわゆる E C C 符号が付されている。この E C C 符号は、セクタ 1 1 1 a から読み出された記憶情報についての誤り検出と、その読み出された情報から誤りが検出された場合の情報修正に用いられる。

【 0 0 8 2 】

E C C 処理回路 1 0 3 は、この E C C 符号を用いて誤り検出と情報修正とを行う回路で

50

ある。

【 0 0 8 3 】

R O M 1 0 4 には、制御 I C 1 0 2 が実行する各種処理を記述した各種プログラムと、各 H D D 1 0 0 についてのアドレス変換情報が格納されている。制御 I C 1 0 2 は、この R O M 1 0 4 内のプログラムに従って処理を実行する。

【 0 0 8 4 】

R A M 1 0 5 には、R O M 1 0 4 中のプログラムが展開されると共に、処理の実行で使われる各種パラメータが格納される。

【 0 0 8 5 】

半導体メモリ 1 0 6 には、後述の交代処理における交代元のセクタ 1 1 1 a の物理アドレスと、そのセクタ 1 1 1 a に割り当てられた交代先のセクタ 1 1 1 a の物理アドレスとを一対一に対応付けた交代対応表が格納される。この交代対応表は、交代処理が実行される度に更新される。

10

【 0 0 8 6 】

尚、図 2 に示す R A I D コントローラ 7 0 の半導体メモリ 7 4 には、この R A I D システム 2 に搭載されている複数の H D D 1 0 0 それぞれについて交代対応表が格納されている。また、本実施形態では、半導体メモリ 7 4 内の各 H D D 1 0 0 の交代対応表には、その H D D 1 0 0 の識別番号が付されている。この R A I D コントローラ 7 0 の半導体メモリ 7 4 内の交代対応表も、交代処理が実行される度に更新される。

【 0 0 8 7 】

20

図 4 は、図 2 に示す R A I D システムに搭載されている H D D を、その H D D の機能に注目して示す機能ブロック図である。

【 0 0 8 8 】

この図 4 に示すように、H D D 1 0 0 は、機能ブロックとして、アクセス部 1 2 1 と、判定部 1 2 2 と、繰返し部 1 2 3 と、第 1 出力部 1 2 4 と、第 2 出力部 1 2 5 とを備えている。

【 0 0 8 9 】

アクセス部 1 2 1 は、R A I D コントローラ 7 0 から先頭 L B A とブロック数とで指定されたアクセス範囲の各セクタ 1 1 1 a に情報を読み書きするものである。ここで、このアクセス部 1 2 1 では、先頭 L B A が物理アドレスに変換される。さらに、アクセス部 1 2 1 は、図 3 に示す R O M 1 0 4 内の交代対応表に、アクセス範囲内の物理アドレスが存在しているか否かを調べる。

30

【 0 0 9 0 】

そして、アクセス部 1 2 1 は、その変換された物理アドレスのセクタ 1 1 1 a から順次にアクセスして情報を読み書きする。ただし、交代対応表に、アクセス範囲内の物理アドレスが存在していた場合には、アクセス部 1 2 1 は、その物理アドレスに対応している交代先の物理アドレスのセクタ 1 1 1 a にアクセスして情報を読み書きする。

【 0 0 9 1 】

このアクセス部 1 2 1 は、R A I D コントローラ 7 0 から書込み指示があった場合には、アクセス範囲の全てのセクタ 1 1 1 a に情報を一旦書き込んだ後に、書込みが完了したか否かの確認のために、この書込み後の各セクタ 1 1 1 a から記憶情報を読み出す。

40

【 0 0 9 2 】

また、アクセス部 1 2 1 は、R A I D コントローラ 7 0 から読出し指示があった場合には、アクセス範囲の各セクタ 1 1 1 a から記憶情報を読み出す。

【 0 0 9 3 】

いずれの場合でも、アクセス部 1 2 1 は、先頭 L B A を変換した物理アドレスのセクタ 1 1 1 a から順次に記憶情報を読み出す。そして、アクセス部 1 2 1 は、各セクタ 1 1 1 a から記憶情報を読み出す度に、その読み出した記憶情報を、その記憶情報を読み出したセクタ 1 1 1 a の物理アドレスと共に判定部 1 2 2 に渡す。

【 0 0 9 4 】

50

また、このアクセス部 1 2 1 は、R A I D コントローラ 7 0 から指定されたアクセス範囲におけるブロック数に基づいて、後述のリトライ時間の合計時間と比較される制限時間を決定する。アクセス部 1 2 1 は、その決定した制限時間を、図 3 に示す R A M 1 0 5 に確保された制限時間格納エリアに格納する。

【 0 0 9 5 】

このアクセス部 1 2 1 は、ハードウェアとしては、図 3 に示す制御 I C 1 0 2 とヘッド 1 0 1 とがその役割を担っている。また、このアクセス部 1 2 1 が、記憶媒体上のアクセス範囲にアクセスすることで情報を読み出す読出部の一例に相当する。

【 0 0 9 6 】

判定部 1 2 2 は、アクセス部 1 2 1 から渡された各セクタ 1 1 1 a の記憶情報から正確な情報が得られたか否かを、その記憶情報に付されている E C C 符号を使った誤り検出によって次のように判定する。

10

【 0 0 9 7 】

判定部 1 2 2 は、アクセス部 1 2 1 から渡された記憶情報に対する誤り検出を E C C 符号を使って行う。そして、その記憶情報から誤りが検出された場合には、判定部 1 2 2 は、その記憶情報を E C C 符号を使って修正する。

【 0 0 9 8 】

判定部 1 2 2 は、記憶情報から誤りが検出されなかった場合、あるいは、誤りが検出され且つ修正ができた場合、アクセス部 1 2 1 に渡された記憶情報から正確な情報が得られたと判定する。判定部 1 2 2 は、正確な情報が得られたと判定した場合、アクセス部 1 2 1 から渡された記憶情報あるは修正後の記憶情報を繰返し指示部 1 2 3 に渡す。

20

【 0 0 9 9 】

一方、判定部 1 2 2 は、誤りが検出され且つ修正ができなかった場合には、アクセス部 1 2 1 から渡された記憶情報から正確な情報が得られなかったと判定する。判定部 1 2 2 は、正確な情報が得られなかったと判定した場合、その判定結果を繰返し指示部 1 2 3 に渡す。また、このときには、判定部 1 2 2 は、正確な情報が得られなかった記憶情報が読み出されたセクタ 1 1 1 a の物理アドレス（読出し不可アドレス）をその判定結果に付して繰返し指示部 1 2 3 に渡す。

【 0 1 0 0 】

この判定部 1 2 2 は、ハードウェアとしては、図 3 に示す制御 I C 1 0 2 と E C C 処理回路 1 0 3 とがその役割を担っている。また、この判定部 1 2 2 が、読出部により読み出された情報から正確な情報が得られたか否かを判定する判定部の一例に相当する。

30

【 0 1 0 1 】

繰返し指示部 1 2 3 は、正確な情報が得られなかった旨を示す判定結果と読出し不可アドレスとの組、あるいは、正確な情報が得られたと判定された記憶情報が判定部 1 2 2 から渡される度に、次のような処理を行う。

【 0 1 0 2 】

即ち、判定部から記憶情報が渡された場合、繰返し指示部 1 2 3 は、その記憶情報を第 1 出力部 1 2 4 に渡す。

【 0 1 0 3 】

一方、正確な情報が得られなかった旨を示す判定結果と読出し不可アドレスとの組が判定部 1 2 2 から渡された場合、繰返し指示部 1 2 3 は、その読出し不可アドレスのセクタ 1 1 1 a についての再読出しを指示する。

40

【 0 1 0 4 】

そして、アクセス部 1 2 1 は、図 3 に示すヘッド 1 0 1 が、再読出しが指示されたセクタ 1 1 1 a に再度近接したタイミングで再度の読出しを実行する。そして、再度の読出しで読み出された情報は、再び、判定部 1 2 2 での判定に供される。この再度の読出し、および再判定を、以下、リトライ処理と呼ぶ。

【 0 1 0 5 】

このリトライ処理は、読出し不可アドレスのセクタ 1 1 1 a から読み出され、判定部 1

50

2 2 で正確な情報が得られたと判定された記憶情報が、判定部 1 2 2 から、繰返し指示部 1 2 3 に渡されるまで繰り返される。そして、1 回もしくは複数回のリトライ処理を経て、その読出し不可アドレスのセクタ 1 1 1 a の記憶情報が判定部 1 2 2 から渡されると、繰返し指示部 1 2 3 は、その記憶情報を第 1 出力部 1 2 4 に渡す。

【 0 1 0 6 】

また、繰返し指示部 1 2 3 は、アクセス部 1 2 1 に対し、再読出しを指示した後、判定部 1 2 2 から、記憶情報、あるいは、読出し不可アドレスと判定結果との組が渡されるまでの経過時間（リトライ時間）を不図示のタイマを使って計時する。

【 0 1 0 7 】

図 3 に示す R A M 1 0 5 には、リトライ時間の保管エリアが確保されている。繰返し指示部 1 2 3 は、リトライ時間を計時する度に、その保管エリアに保管されていたリトライ時間に、今回計時したリトライ時間を加算する。これにより、リトライ処理が複数回繰り返されたときのリトライ時間の合計時間が算出される。

10

【 0 1 0 8 】

そして、繰返し指示部 1 2 3 は、図 3 に示す R A M 1 0 5 に確保された上記の制限時間格納エリアに格納されている制限時間と、この合計時間とを比較する。合計時間が制限時間よりも少ない場合には、繰返し指示部 1 2 3 は、上記の保管エリアにその合計時間を上書きする。一方、合計時間が制限時間以上である場合には、その旨を第 1 出力部 1 2 4 に通知する。

【 0 1 0 9 】

20

この繰返し指示部 1 2 3 は、ハードウェアとしては、図 3 に示す制御 I C 1 0 2 がその役割を担っている。また、この繰返し指示部 1 2 3 が、正確な情報が得られなかったと判定部により判定された場合に読出部に、アクセス範囲からの記憶情報の読出しを繰り返させる繰返し指示部の一例に相当する。

【 0 1 1 0 】

第 1 出力部 1 2 4 には、上述したように、繰返し指示部 1 2 3 から、正確な情報が得られたと判定部 1 2 2 で判定された、アクセス範囲の各セクタ 1 1 1 a の記憶情報あるいは修正後の記憶情報が渡される。

【 0 1 1 1 】

ここで、上述したように、アクセス部 1 2 1 が実行する読出しには、R A I D コントローラ 7 0 からの読出し指示に応じて行われるものと、情報が書き込まれた後に、その書込みが完了したか否かの確認のために行われるものとがある。

30

【 0 1 1 2 】

第 1 出力部 1 2 4 は、R A I D コントローラ 7 0 からの指示が読出し指示である場合には、繰返し指示部 1 2 3 から渡された記憶情報を、R A I D コントローラ 7 0 に向けて出力する。さらに、アクセス範囲の全ての物理アドレスについて記憶情報の出力が終了した段階で、第 1 出力部 1 2 4 は、読出しが完了した旨を示す完了通知情報を R A I D コントローラ 7 0 に向けて出力する。

【 0 1 1 3 】

一方、第 1 出力部 1 2 4 は、R A I D コントローラ 7 0 からの指示が書込み指示である場合には、アクセス範囲の全ての物理アドレスについて記憶情報が渡された段階で、完了通知情報を R A I D コントローラ 7 0 に向けて出力する。この完了通知情報は、書込みが完了した旨を示している。

40

【 0 1 1 4 】

また、完了通知情報を出力すると、第 1 出力部 1 2 4 は、R A M 1 0 5 における制限時間格納エリアに格納されている制限時間、および、リトライ時間の保管エリアに格納されている合計時間を消去する。

【 0 1 1 5 】

ここで、上述したように、繰返し指示部 1 2 3 は、リトライ処理が行われる度に算出されるリトライ時間の合計時間が制限時間以上である場合に、第 1 出力部 1 2 4 にその旨を

50

通知する。

【 0 1 1 6 】

第 1 出力部 1 2 4 は、繰返し指示部 1 2 3 から、リトライ時間の合計時間が制限時間以上である旨が通知されてきた場合、次のようなエラー情報を R A I D コントローラ 7 0 に向けて出力する。エラー情報は、アクセス範囲からの情報の読出しに失敗したこと、あるいは、書込み後の確認のための読出しに失敗したことを示す情報である。ここで、書込み後の確認のための読出しに失敗したことは、情報が正常に書き込めなかったこと、即ち、情報の書込みに失敗したことを意味している。

【 0 1 1 7 】

また、第 1 出力部 1 2 4 は、エラー情報を出力したときには、アクセス部 1 2 1、判定部 1 2 2、および繰返し指示部 1 2 3 それぞれに対し、それ以降の動作を停止させる。

10

【 0 1 1 8 】

また、このときにも、第 1 出力部 1 2 4 は、図 3 に示す R A M 1 0 5 に格納されている制限時間および合計時間を消去する。

【 0 1 1 9 】

この第 1 出力部 1 2 4 は、ハードウェアとしては、図 3 に示す制御 I C 1 0 2 がその役割を担っている。また、この第 1 出力部 1 2 4 が、正確な情報が得られた場合には情報およびアクセス完了通知の少なくとも一方を出力し、読出しが限度まで繰り返されても、正確な情報が得られなかった場合にはエラー情報を出力する第 1 出力部の一例に相当する。

【 0 1 2 0 】

20

第 2 出力部 1 2 5 は、R A I D コントローラ 7 0 から、後述の生データ読出し指示を受けた場合に、アクセス部 1 2 1 に対し次のような指示を行う。

【 0 1 2 1 】

この場合、第 2 出力部 1 2 5 は、アクセス部 1 2 1 に対し、生データ読出し指示において後述するように指定されたアクセス範囲の各セクタ 1 1 1 a から記憶情報を読み出して、判定部 1 2 2 には渡さずに第 2 出力部 1 2 5 に渡すように指示する。第 2 出力部 1 2 5 は、この指示に応じてアクセス部 1 2 1 から渡された記憶情報を R A I D コントローラ 7 0 に向けて出力する。

【 0 1 2 2 】

この第 2 出力部 1 2 5 は、ハードウェアとしては、図 3 に示す制御 I C 1 0 2 がその役割を担っている。また、この第 2 出力部 1 2 5 が、特定の指示を受けた場合に、読出部により読み出された記憶情報をそのまま出力する第 2 出力部の一例に相当する。

30

【 0 1 2 3 】

以上に説明した H D D 1 0 0 が、記憶媒体と、読出部と、判定部と、繰返し指示部と、第 1 出力部と、第 2 出力部とを備えた記憶装置の一例に相当する。

【 0 1 2 4 】

以上で、図 4 を参照した H D D の機能ブロックについての説明を終了し、次に、図 2 に示す R A I D コントローラ 7 0 について機能に注目して説明する。

【 0 1 2 5 】

図 5 は、図 2 に示す R A I D システムに搭載されている R A I D コントローラを、その R A I D コントローラの機能に注目して示す機能ブロック図である。

40

【 0 1 2 6 】

この図 5 に示すように、R A I D コントローラ 7 0 は、機能ブロックとして、リカバリリミット処理部 7 5 と、媒体データ格納部 7 6 と、データ正当性分析部 7 7 と、アドレス保持部 7 8 とを備えている。リカバリリミット処理部 7 5 およびデータ正当性分析部 7 7 は、ハードウェアとしては、図 2 に示す制御 I C 7 1 がその役割を担っている。また、媒体データ格納部 7 6 およびアドレス保持部 7 8 は、各々、図 2 に示す R A M 7 3 内における予め決められた領域がその役割を担っている。

【 0 1 2 7 】

以下、図 5 に示すこれらの機能ブロックの詳細について、R A I D コントローラ 7 0 で

50

実行される処理と併せて説明する。

【0128】

図6は、RAIDコントローラで実行される処理の流れを示すフローチャートである。

【0129】

このフローチャートが示す処理は、このRAIDシステム2に、上位装置から、ライトコマンドあるいはリードコマンドが送信されて来るとスタートする。これらのコマンドでは、アクセス範囲が上述の先頭LBAとブロック数とで指定される。また、ライトコマンドが送信されて来たときには、そのライトコマンドに続いて書込み情報が上位装置から送信される。

【0130】

上位装置から送信されて来るコマンドは、リカバリリミット処理部75で受信される(ステップS101)。そして、リカバリリミット処理部75は、RAIDシステム2内の複数のHDD100のうち、コマンドで指定されているアクセス範囲が属するHDD100に対し、そのコマンドに応じた指示を行う。

【0131】

ここで、指示対象のHDD100の把握は、図2に示すROM72に、各HDD100の識別番号が付されて格納されているアドレス変換情報を参照することで行われる。LBAは、各セクタ111aに割り当てられた、RAIDシステム2内の複数のHDD100に亘る通し番号である。このため、各HDD100内のセクタ111aに割り当てられるLBAの範囲が、HDD100の相互間で異なっている。各HDD100のアドレス変換情報は、LBAから、そのHDD100内のセクタ111aの物理アドレスへの変換関係を表したものである。従って、各HDD100のアドレス変換情報には、そのHDD100内のセクタ111aの物理アドレスに変換されるべきLBAの範囲が対応している。

【0132】

リカバリリミット処理部75は、コマンドで指定されているアクセス範囲を含んだLBAの範囲に対応したアドレス変換情報に付されている識別番号のHDD100を、指示対象のHDD100として把握する。

【0133】

そして、上位装置からのコマンドがライトコマンドの場合には、リカバリリミット処理部75は、そのHDD100に対し情報の書込みを指示する。また、上位装置からのコマンドがリードコマンドの場合には、リカバリリミット処理部75は、そのHDD100に対し情報の読出しを指示する。

【0134】

このステップS101の処理を実行しているリカバリリミット処理部75が、記憶装置にアクセス範囲を指定する指定部の一例に相当する。

【0135】

この後、リカバリリミット処理部75は、そのHDD100から、上記の完了通知情報、読出し指示に応じて読み出された記憶情報、あるいはエラー情報が出力されて来るまで待機する(ステップS102)。

【0136】

HDD100から、完了通知情報や読出し指示に応じて読み出された記憶情報(読出し情報)が出力された場合には(ステップS102におけるNo判定)、リカバリリミット処理部75は、次のステップS103の処理を実行する。

【0137】

ステップS103では、上位装置からのコマンドがライトコマンドであって、HDD100から完了通知情報が出力されて来た場合には、リカバリリミット処理部75が、コマンドに応じた処理の完了を示すコマンド完了通知情報を上位装置に向けて出力する。

【0138】

また、ステップS103では、コマンドがリードコマンドであって、HDD100から完了通知情報と読出し情報が出力されて来た場合には、リカバリリミット処理部75が、

10

20

30

40

50

コマンド完了通知情報と読出し情報とを上位装置に向けて出力する。

【0139】

ステップS103の処理が実行された場合には、図6のフローチャートが表す処理はそのまま終了する。

【0140】

一方、HDD100からエラー情報が出力された場合には（ステップS102におけるYes判定）、リカバリリミット処理部75は、後述のステップS104以降の処理を実行する。

【0141】

ここで、ステップS104以降の処理について説明する前に、HDD100がエラー情報を出力する状況について説明する。

【0142】

図7は、HDDがエラー情報を出力する状況の一例を模式的に示す図である。

【0143】

この図7には、HDD100に対して20個のブロック数のアクセス範囲A1が指定されて読出し指示が行われたときの、このHDD100で実行される処理が示されている。

【0144】

図4を参照して説明したように、HDD100では、各セクタ111aから情報が読み出されるときには、ECCによる修正も含めて正確な情報が得られるまで、繰返し指示部123の指示によってリトライ処理が繰り返される。図7の例では、先頭（即ち「1」番目）から「4」番目まで、「6」番目、「9」番目、および「11」番目から「17」番目までの各セクタ111aについては、1回の読出しで正確な情報が得られている。

【0145】

一方、先頭から「5」番目のセクタ111aでは、1回目の読出しで正確な情報が得られず、その後リトライ処理が4回繰り返されている。そして、4回目のリトライ処理（読出しとしては5回目）で正確な情報が得られている。「7」番目のセクタ111aでは、10回目のリトライ処理（読出しとしては11回目）で正確な情報が得られている。「8」番目のセクタ111aでは、5回目のリトライ処理（読出しとしては6回目）で正確な情報が得られている。「10」番目のセクタ111aでは、2回目のリトライ処理（読出しとしては3回目）で正確な情報が得られている。

【0146】

ここで、図4の繰返し指示部123では、各セクタ111aについて計時されたリトライ時間が、リトライ処理が実行された全セクタ111aに亘って合計されて、リトライ時間の合計時間が算出される。そして、繰返し指示部123では、その合計時間が、図3に示すRAM105中の制限時間と比較される。この合計時間が制限時間以上である場合に、図4の第1出力部124がエラー情報を出力する。図7の例では、「18」番目のセクタ111aにおいて、1回目の読出しで正確な情報が得られずにリトライ処理が行われたところ、そのリトライ処理の処理時間を含めた合計時間が制限時間以上となっている。そして、この合計時間が制限時間以上となっていることを受けて、第1出力部124がエラー情報を出力する。

【0147】

この図7に一例として示されている状況では、エラー情報出力時点でリトライ処理が行われていたのは、「18」番目のセクタ111aであって、その時に行われていたリトライ処理は1回目のリトライ処理である。一方、この「18」番目のセクタ111a以前に読出しが行われたセクタ111aの中には、例えば「7」番目のセクタ111aのように多数回に及びリトライ処理が行われたものがある。このように、リトライ処理が多数回に亘って行われるセクタ111aには、傷等の欠陥が存在している可能性が高い。

【0148】

つまり、図7に示されている状況は、エラー情報出力時点でリトライ処理が行われていたセクタ111aとは別に、欠陥を有する可能性の高いセクタ111aが存在していると

10

20

30

40

50

いう状況である。

【0149】

以下、この図7に示されている状況を適宜に参照しながら、図6のフローチャートについての説明を続ける。

【0150】

上述したように、HDD100からエラー情報が出力された場合には(ステップS102におけるYes判定)、リカバリリミット処理部75はステップS104以降の処理を実行する。

【0151】

ステップS104では、リカバリリミット処理部75が、そのエラー情報を出力したHDD100に対し、次のような生データ読出し指示を行う。

【0152】

生データ読出し指示とは、ライトコマンドあるいはリードコマンドで指定されたアクセス範囲について、ECCによる誤り検出、修正、およびリトライ処理を行わずに、各セクタ111aにつき1回だけ記憶情報を読み出す行うように指示することである。

【0153】

この生データ読出し指示を受けたHDD100では、上述したように、第2出力部125が、アクセス範囲の各セクタ111aにつき1回ずつ読み出された記憶情報を順次に出力する。

【0154】

リカバリリミット処理部75は、そのように第2出力部125から順次に出力されて来る各セクタ111aの記憶情報を、媒体データ格納部76に格納する(ステップS105)。この媒体データ格納部76には、各セクタ111aの記憶情報が、そのセクタ111aのLBAと対応付けられて格納される。

【0155】

ここで、このステップS105で媒体データ格納部76に格納される記憶情報は、上記のようにECCによる誤り検出、修正、およびリトライ処理が行われずに読み出された情報である。このため、媒体データ格納部76に格納されたこれらの記憶情報の中には、誤りを含んだ記憶情報が存在している可能性が高い。

【0156】

例えば、上記の図7に示す状況では、リトライ処理が実行された「5」番目、「7」番目、「8」番目、「10」番目、「18」番目それぞれのセクタ111aからステップS105で読み出された記憶情報が誤りを含んでいる可能性が高い。

【0157】

媒体データ格納部76に記憶情報が格納されると、次に、データ正当性分析部77が、その格納された各記憶情報について、次のステップS106の処理を行う。

【0158】

ステップS106では、データ正当性分析部77は、まず、アクセス範囲における先頭LBAから順次に、図2に示すROM72内のアドレス変換情報を使って物理アドレスに変換する。そして、データ正当性分析部77は、変換後の各物理アドレスが、図2に示す半導体メモリ74内の交代対応表に交代元として記載されているか否かを判定する。

【0159】

さらに、ステップS106では、データ正当性分析部77は、変換後の物理アドレスが交代元として記載されていなかった場合に、その変換前のLBAのセクタ111aの記憶情報について次のように誤り検出を行う。データ正当性分析部77は、媒体データ格納部76から検出対象の記憶情報を読み出す。そして、データ正当性分析部77は、その読み出した記憶情報に付されているECC符号を使って、その記憶情報について誤り検出を行う。

【0160】

データ正当性分析部77は、この誤り検出の結果から、その検出対象の記憶情報が、誤

10

20

30

40

50

りが検出されない情報（即ち、正当な情報）であるか否かを判定する（ステップS107）。また、ステップS107では、データ正当性分析部77は、変換後の物理アドレスが交代元として交代対応表に記載されていたLBAが対応付けられている記憶情報については、上記の誤り検出を行わずに正当な情報ではないと判定する。

【0161】

記憶情報が正当な情報ではない場合（ステップS107におけるNo判定）、データ正当性分析部77は、媒体データ格納部76においてその記憶情報に対応付けられているLBAをアドレス保持部78に格納する（ステップS108）。その後、データ正当性分析部77は、その記憶情報に対応付けられているLBAが、アクセス範囲における最後のLBAであるか否かを判定する（ステップS109）。

10

【0162】

一方、記憶情報が正当な情報であった場合（ステップS107におけるYes判定）、データ正当性分析部77は、上記のステップS108の処理を省略してステップS109の判定を行う。

【0163】

記憶情報に対応付けられているLBAが、アクセス範囲における最後のLBAではなかった場合には（ステップS109におけるNo判定）、データ正当性分析部77は、ステップS106以降の処理を繰り返す。この繰り返しは、記憶情報に対応付けられているLBAが最後のLBAであると判定（ステップS109におけるYes判定）されるまで繰り返される。

20

【0164】

ステップS106～S109の処理により、アクセス範囲の各LBAのうち、記憶情報が正当な情報ではないと判定されたLBAが、アドレス保持部78に格納される。

【0165】

図8は、アクセス範囲の各LBAについての、記憶情報が正当な情報であるか否かの判定結果の一例を模式的に示す図である。

【0166】

この図8に示す例は、図7に示す状況で出力されたエラー情報を受けた生データ読み出し指示に応じて読み出された記憶情報についての判定結果である。

【0167】

この図8の例では、図7に示す状況でリトライ処理が1回でも行われた記憶情報が、全て、正当な情報ではないと判定されている。即ち、今回のアクセス範囲A1において、先頭から「5」番目、「7」番目、「8」番目、「10」番目、「18」番目それぞれのセクタ111aの記憶情報が正当な情報ではないと判定されている。

30

【0168】

この図8の例では、「5」番目、「7」番目、「8」番目、「10」番目、「18」番目の5個のセクタ111aそれぞれのLBAが、記憶情報が正当な情報ではないと判定されたLBA（不良LBA）としてアドレス保持部78に格納される。

【0169】

尚、ステップS108でアドレス保持部78に格納されるLBAには、そのLBAに対応する記憶情報から誤りが検出されたものと、そのLBAを変換した物理アドレスが交代対応表に交代元として記載されているものとの2種類がある。上記の図8の例では、アドレス保持部78に格納されるLBAは、記憶情報から誤りが検出されたLBAのみとなっている。

40

【0170】

図6のステップS106～S109の処理で、このように不良LBAがアドレス保持部78に格納されると、次に、リカバリリミット処理部75およびデータ正当性分析部77は以下のステップS110の処理を行う。

【0171】

ステップS110では、リカバリリミット処理部75は、磁気ディスク110における

50

次のような2つの隣接範囲を求める。隣接範囲は、今回のアクセス範囲のセクタ111aが属するトラック111を挟んで、このトラックに隣接する2本の隣接トラック111それぞれ上で、今回のアクセス範囲に隣接する範囲である。

【0172】

これら2つの隣接範囲は、次のように求められる。

【0173】

まず、リカバリリミット処理部75は、今回のアクセス範囲の先頭LBAを、図2に示すROM72内のアドレス変換情報を使って物理アドレスに変換する。上述したように、物理アドレスは、シリンダ番号とヘッド番号とセクタ番号とで、各セクタ111aの位置を示すCHS形式で記述されたアドレスである。

10

【0174】

リカバリリミット処理部75は、先頭LBAから変換した物理アドレス（先頭物理アドレス）におけるシリンダ番号への「1」の加算と、先頭物理アドレスにおけるシリンダ番号の「1」の減算とを行う。先頭物理アドレスのシリンダ番号に「1」を加算したシリンダ番号は、アクセス範囲のセクタ111aが属するトラック111の外周側に隣接するトラック111に対応したシリンダ番号である。また、先頭物理アドレスのシリンダ番号から「1」を減算したシリンダ番号は、アクセス範囲のセクタ111aが属するトラック111の内周側に隣接するトラック111に対応したシリンダ番号である。

【0175】

「1」を加算したシリンダ番号と、先頭物理アドレスにおけるヘッド番号とセクタ番号とで表される物理アドレスが、外周側の隣接トラック上で、今回のアクセス範囲に隣接する範囲の先頭物理アドレスとなる。そして、この先頭物理アドレスと、今回のアクセス範囲のブロック数とで表される範囲が、外周側の隣接トラック上における隣接範囲（外周側隣接範囲）となる。

20

【0176】

また、「1」を減算したシリンダ番号と、先頭物理アドレスにおけるヘッド番号とセクタ番号とで表される物理アドレスが、内周側の隣接トラック上で、今回のアクセス範囲に隣接する範囲の先頭物理アドレスとなる。そして、この先頭物理アドレスと、今回のアクセス範囲のブロック数とで表される範囲が、内周側の隣接トラック上における隣接範囲（内周側隣接範囲）となる。

30

【0177】

ステップS110では、リカバリリミット処理部75が、このようにCHS形式で求めた2つの隣接範囲それぞれを、CHS形式のままHDD100に順次に指定して、各隣接範囲からの記憶情報の読出しを指示する。このときの読出しについても、ステップS104での指示と同様に、ECCによる誤り検出、修正、およびリトライ処理を行わずに、各セクタ111aにつき1回だけ読出す行うように指示される。

【0178】

内周側、外周側それぞれの隣接範囲から読み出されてHDD100から出力されてきた記憶情報が、リカバリリミット処理部75によって、媒体データ格納部76に格納される。このとき、記憶情報は、その記憶情報が読み出されたセクタ111aの物理アドレスに対応付けられて格納される。

40

【0179】

そして、ステップS110では、データ正当性分析部77によって、ステップS106～S109の処理と同様の処理が行われる。これにより、内周側、外周側それぞれの隣接範囲において、正当な情報ではない記憶情報が読み出されたセクタ111aの物理アドレス（不良物理アドレス）が、アドレス保持部78に格納される。

【0180】

尚、このステップS110でも、誤り検出は、各記憶情報のうち、その記憶情報に対応する物理アドレスが交代対応表に交代元として記載されていないものについてのみ行われる。そして、交代対応表に交代元として記載されている物理アドレスについては、その物

50

理アドレスに対応付けられている記憶情報に対する誤り検出を行わずに、その物理アドレスが、不良物理アドレスとしてアドレス保持部 7 8 に格納される。

【 0 1 8 1 】

上記のステップ S 1 0 4 の処理とステップ S 1 1 0 の処理とを実行しているリカバリリミット処理部 7 5 が、次のような指示付与部の一例に相当する。指示付与部は、記憶装置から出力された情報がエラー情報である場合に、第 1 アクセス範囲と第 2 アクセス範囲とを指定するとともに、第 2 出力部に対して特定の指示を与えるものである。第 1 アクセス範囲は、情報読出しが試行された記憶区画である試行区画を含んだ範囲である。第 2 アクセス範囲は、試行区画のトラックに隣接するトラック上の、その試行区画に隣接する記憶区画である隣接区画を含んだ範囲である。

10

【 0 1 8 2 】

本実施形態では、ステップ S 1 0 4 の生データ読出し指示で指定されるアクセス範囲が、第 1 アクセス範囲の一例に相当する。また、ステップ S 1 1 0 で指定される内周側、外周側それぞれの隣接範囲が、各々第 2 アクセス範囲の一例に相当する。

【 0 1 8 3 】

尚、本実施形態では、エラー情報は、単に読出しが失敗したことを示す情報となっている。ここで、このエラー情報に、例えば、上記のリトライ処理の合計時間が制限時間以上となったときのセクタ 1 1 1 a の L B A を含ませても良い。そして、第 1 アクセス範囲は、アクセス範囲の先頭 L B A と、そのエラー情報に含まれている L B A のセクタ 1 1 1 a までのブロック数で表される範囲であっても良い。

20

【 0 1 8 4 】

また、上記の第 2 アクセス範囲も、アクセス範囲の先頭 L B A を変換した物理アドレスから、このようなエラー情報に含まれている L B A のセクタ 1 1 1 a までのブロック数で表される範囲であっても良い。

【 0 1 8 5 】

さらに、第 2 アクセス範囲については、正当な情報ではないと判定された記憶情報について情報読出しが試行されたセクタ 1 1 1 a に隣接するセクタ 1 1 1 a を含む範囲であれば、隣接トラック上のどのような範囲であっても良い。

【 0 1 8 6 】

ステップ S 1 1 0 の処理が終了すると、リカバリリミット処理部 7 5 が、次のようなマッピング処理を行う (ステップ S 1 1 1)。

30

【 0 1 8 7 】

このマッピング処理は、内周側、外周側それぞれの隣接範囲内の不良物理アドレスの中から、今回のアクセス範囲内の不良 L B A のセクタ 1 1 1 a に隣接するセクタ 1 1 1 a (隣接不良セクタ) の物理アドレスを選び出す処理である。

【 0 1 8 8 】

このマッピング処理では、リカバリリミット処理部 7 5 は、まず、ステップ S 1 0 8 でアドレス保持部 7 8 に格納された不良 L B A を物理アドレスに変換する。リカバリリミット処理部 7 5 は、この変換も、図 2 に示す R O M 7 2 内のアドレス変換情報を使って行う。

40

【 0 1 8 9 】

そして、リカバリリミット処理部 7 5 は、内周側、外周側それぞれの隣接範囲内の不良物理アドレスの中から、不良 L B A を変換した物理アドレスのヘッド番号およびセクタ番号と同じヘッド番号とセクタ番号とを有するものを選び出す。このように選び出された不良物理アドレスが、内周側、外周側それぞれの隣接範囲内における隣接不良セクタの物理アドレスに相当する。

【 0 1 9 0 】

図 9 は、マッピング処理を模式的に示す図である。

【 0 1 9 1 】

この図 9 には、上記の図 7 に示す状況下でのマッピング処理が示されている。そして、

50

この図9には、図7および図8に示されたアクセス範囲A1と、そのアクセス範囲A1に隣接する内周側隣接範囲A2と外周側隣接範囲A3とが示されている。

【0192】

この図9のパート(A)には、アクセス範囲A1における不良LBAのセクタA1__1と、内周側隣接範囲A2における不良物理アドレスのセクタA2__1と、外周側隣接範囲A3における不良物理アドレスのセクタA3__1とが示されている。

【0193】

不良LBAのセクタA1__1は、アクセス範囲A1の先頭から「5」番目、「7」番目、「8」番目、「10」番目、「18」番目の5個のセクタである。

【0194】

内周側隣接範囲A2における不良物理アドレスのセクタA2__1は、内周側隣接範囲A2の先頭から「6」番目、「7」番目、「8」番目、「10」番目の4個のセクタである。

【0195】

外周側隣接範囲A3における不良物理アドレスのセクタA3__1は、外周側隣接範囲A3の先頭(即ち、「1」番目)、「6」番目、「7」番目、「8」番目、「9」番目、「18」番目の6個のセクタである。

【0196】

上述したように、上記のマッピング処理では、内周側、外周側それぞれの隣接範囲内の不良物理アドレスの中から、アクセス範囲内の不良LBAのセクタ111aに隣接する隣接不良セクタの物理アドレスが選出される。

【0197】

図9の例では、不良物理アドレスのセクタA2__1, A3__1のうち、内周側、外周側それぞれの隣接範囲での番数が、アクセス範囲での不良LBAのセクタA1__1の番数と同じものが、内周側、外周側それぞれの隣接不良セクタとなる。内周側隣接範囲A2における不良物理アドレスのセクタA2__1のうち、隣接不良セクタは「7」番目、「8」番目、「10」番目のセクタである。また、外周側隣接範囲A3における不良物理アドレスのセクタA3__1のうち、隣接不良セクタは「7」番目、「8」番目、「18」番目のセクタ111aである。これらの隣接不良セクタは、物理アドレスにおけるシリンダ番号以外のヘッド番号とセクタ番号は、不良LBAから変換された物理アドレスのヘッド番号とセクタ番号と同じである。

【0198】

そこで、上記のマッピング処理では、内周側、外周側それぞれの隣接範囲内の不良物理アドレスの中から、不良LBAから変換した物理アドレスのヘッド番号およびセクタ番号と同じヘッド番号とセクタ番号とを有するものが選出される。これにより、内周側、外周側それぞれの隣接範囲内で、不良LBAのセクタ111aに隣接する隣接不良セクタの物理アドレスが選出される。

【0199】

図6のフローチャートにおけるステップS111のマッピング処理では、リカバリリミット処理部75は、アドレス保持部78に格納されている不良物理アドレスのうち、このように選出された隣接不良セクタの物理アドレス以外の不良物理アドレスを消去する。

【0200】

次に、リカバリリミット処理部75は、以下に説明する交代処理の処理対象となるセクタ111aの交代対象の物理アドレスを決定する(ステップS112)。

【0201】

交代処理は、交代対象の物理アドレスのセクタ111aに割り当てられていたLBAを、そのセクタ111aに替えて、上記の予備領域中のセクタ111aに割り当て直す処理である。つまり、この交代処理では、交代対象の物理アドレスのセクタ111aが、そのセクタ111aが属するHDD100における複数のセクタ111aのうちの予備領域中のセクタ111aに割り当てられる。

10

20

30

40

50

【 0 2 0 2 】

ステップ S 1 1 2 では、まず、隣接不良セクタの中から、不良 L B A のセクタ 1 1 1 a を内周側と外周側から挟む隣接不良セクタの組が探索される。

【 0 2 0 3 】

図 9 のパート (B) には、この隣接不良セクタの組の探索の一例が示されている。

【 0 2 0 4 】

上述したように、図 9 の例では、内周側の隣接不良セクタは、内周側隣接範囲 A 2 における「 7 」番目、「 8 」番目、「 1 0 」番目のセクタ 1 1 1 a である。また、外周側の隣接不良セクタは、外周側隣接範囲 A 3 における「 7 」番目、「 8 」番目、「 1 8 」番目のセクタ 1 1 1 a である。そして、今回のアクセス範囲 A 1 における不良 L B A のセクタ 1 1 1 a を内周側と外周側から挟む隣接不良セクタの組が、図 9 のパート (B) に、円形エリア C 内に示されている。即ち、この図 9 のパート (B) には、内周側の「 7 」番目の隣接不良セクタと外周側の「 7 」番目の隣接不良セクタとの組、および、内周側の「 8 」番目の隣接不良セクタと外周側の「 8 」番目の隣接不良セクタとの組が示されている。

10

【 0 2 0 5 】

本実施形態では、このような隣接不良セクタの組が、内周側の隣接不良セクタの物理アドレスと、外周側の隣接不良セクタの物理アドレスとの相互間で、シリンダ番号以外の番号 (ヘッド番号とセクタ番号) が同じとなる組を探索することで行われる。

【 0 2 0 6 】

図 6 のフローチャートにおけるステップ S 1 1 2 では、このように隣接不良セクタの組が探索されると、リカバリリミット処理部 7 5 は、次のような物理アドレスを、交代元の物理アドレスとしてアドレス保持部 7 8 に格納する。

20

【 0 2 0 7 】

即ち、リカバリリミット処理部 7 5 は、探索された組に属する隣接不良セクタの物理アドレスと、今回のアクセス範囲内の不良 L B A から変換された不良物理アドレスとを、交代元の物理アドレスとしてアドレス保持部 7 8 に格納する。

【 0 2 0 8 】

次に、リカバリリミット処理部 7 5 は、ステップ S 1 1 2 でアドレス保持部 7 8 に格納した物理アドレスを対象に、次のように交代処理を行う (ステップ S 1 1 3) 。

【 0 2 0 9 】

リカバリリミット処理部 7 5 は、まず、ステップ S 1 1 2 でアドレス保持部 7 8 に格納した物理アドレスのうち、図 2 の半導体メモリ 7 4 内の交代対応表に交代元として未だ格納されていないものを選び出す。そして、その選び出した物理アドレスを、交代対応表における交代元の物理アドレスの格納欄に、新たな交代元の物理アドレスとして格納する。

30

【 0 2 1 0 】

次に、リカバリリミット処理部 7 5 は、予備領域の物理アドレスの中から、交代対応表に交代先として未だ格納されていないものを、所定の優先順位に従って、上記の新たな交代元の交代先として選び出す。そして、リカバリリミット処理部 7 5 は、その選び出した予備領域の物理アドレスを、交代対応表における、上記の新たな交代元の物理アドレスに対応した交代先の格納欄に格納する。

40

【 0 2 1 1 】

リカバリリミット処理部 7 5 は、以上に説明した交代処理で更新された交代対応表を、その交代処理の対象となった HDD 1 0 0 に送信する。交代処理後の交代対応表を受信した HDD 1 0 0 では、図 3 の制御 IC 1 0 2 が、半導体メモリ 1 0 6 に、その受信した交代対応表を上書きする。これにより、HDD 1 0 0 内の交代対応表が最新の交代対応表に更新される。

【 0 2 1 2 】

また、リカバリリミット処理部 7 5 は、新たに交代先となった予備領域のセクタ 1 1 1 a に、次のような情報を書き込む。

【 0 2 1 3 】

50

上述したように、本実施形態のRAIDシステム2におけるRAID構造は、RAID1である。従って、今回アクセスの対象となったHDD100に記憶されている情報と同じ情報が、そのHDD100とは別のHDDに冗長に記憶されている。

【0214】

新たな交代元となった各物理アドレスのセクタ111aからは正確な情報が得られていないことから、今回アクセスの対象となったHDD100では、交代元の物理アドレスに対応する情報が欠損している可能性が高い。

【0215】

そこで、リカバリリミット処理部75は、交代元の物理アドレスに対応する記憶情報の読出しを、その別のHDD100に指示する。そして、その別のHDD100から出力されて来た記憶情報の、交代元の物理アドレスに交代先として割り当てられたセクタ111aへの書込みを、今回アクセスの対象となったHDD100に指示する。これにより、上記のような情報の欠損が埋められることとなる。

【0216】

以上に説明したステップS111～S113の処理を実行しているリカバリリミット処理部75が、次のような交代部の一例に相当する。この交代部は、情報読出が試行された試行区画および隣接区画の双方について正確な情報が得られない場合の試行区画に替えて、複数の記憶区画のうちの交代区画を割り当てるものである。

【0217】

ステップS113の処理が終了すると、この第2実施形態のRAIDシステム2のRAIDコントローラ70における、図6のフローチャートが表す処理が終了する。

【0218】

このRAIDコントローラ70が、指定部と、指示付与部と、交代部とを備えた制御装置の一例に相当する。

【0219】

以上に説明した第2実施形態の情報記憶システム2では、HDD100からエラー情報が出力されると、今回のアクセス範囲の各セクタ111aと、内周側、外周側それぞれの隣接範囲の各セクタ111aから記憶情報がそのまま読み出される。それぞれの範囲のセクタ111aからそのまま読み出された記憶情報の中には、誤りを有する情報が含まれている。ここで、この誤りの中には、リトライ処理を何度も繰り返させる傷等の欠陥による重篤な誤りや、クロストークやノイズ等による、少ない回数のリトライ処理で修正可能な軽微な誤り等といった、軽重様々なものがある。そして、上記のように、重篤な誤りに繋がる欠陥は、隣のトラックにまで及んでいる可能性が高い。

【0220】

この情報記憶システム2では、今回のアクセス範囲の各セクタ111aと隣接範囲の各セクタ111aとの双方について正確な情報が得られない場合に、そのセクタ111aについて交代処理が行われる。これにより、互いに隣り合った複数のトラック間に亘る欠陥が存在している可能性が高いセクタ111aについて交代処理が行われることとなる。

【0221】

そして、次回に、このセクタ111aを含むアクセス範囲からの情報の読出しが情報記憶システム2において実行されるときには、そのセクタ111aに替えて交代先のセクタ111aにアクセスされることとなる。その結果、上記のようなリトライ処理の繰返しの発生が抑えられる。つまり、この情報記憶システム2では、パフォーマンスの低下が抑えられることとなる。

【0222】

また、この第2実施形態の情報記憶システム2では、図9のステップS110の処理において、隣接範囲として、アクセス範囲が属するトラック111を挟んだ2つの隣接トラックそれぞれ上の、そのアクセス範囲に隣接する2つの隣接範囲が指定される。そして、これら2つの隣接範囲のいずれについても正確な情報が得られないセクタ111aが交代処理の対象とされる。これにより、複数のトラックに亘る重篤な欠陥が存在するセクタ1

10

20

30

40

50

1 1 a が、高い精度で見つけられることとなる。このことは、本件の情報記憶システムに対し、以下の応用形態が好適であることを意味している。

【 0 2 2 3 】

この応用形態では、上記指示付与部が、上記第 2 アクセス範囲として、次のような 2 つの第 2 アクセス範囲を上記記憶装置に指定するものとなっている。2 つの第 2 アクセス範囲は、上記の試行区画のトラックにそれぞれが隣接して相互間にそのトラックを挟んだ 2 つのトラックそれぞれ上の隣接区画を含んだ 2 つの範囲である。また、この応用形態では、上記交代部が、次のような試行区画について上記交代区画を割り当てるものとなっている。この交代部は、上記試行区画と、その試行区画を挟んだ両方の上記隣接区画とのいずれについても正確な情報が得られない場合のその試行区画について上記交代区画を割り当てる。

10

【 0 2 2 4 】

本実施形態において、図 6 のステップ S 1 0 4 の処理とステップ S 1 1 0 の処理とを実行しているリカバリリミット処理部 7 5 は、この応用形態における指示付与部の一例にも相当している。また、図 6 のステップ S 1 1 1 ~ S 1 1 3 の処理を実行しているリカバリリミット処理部 7 5 は、この応用形態における交代部の一例にも相当している。

【 0 2 2 5 】

尚、本実施形態では、上記のように、欠陥が存在するセクタ 1 1 1 a の発見精度を高めるために、隣接範囲として、アクセス範囲が属するトラック 1 1 1 を挟んだ 2 つの隣接トラックそれぞれ上の、そのアクセス範囲に隣接する 2 つの隣接範囲が指定される。しかしながら、本件の情報記憶システムは、この形態に限るものではない。本件の情報記憶システムは、例えば、アクセス範囲が属するトラックに隣接する 1 本の隣接トラック上の、そのアクセス範囲に隣接する 1 つの隣接範囲が指定されるものであっても良い。

20

【 0 2 2 6 】

また、この第 2 実施形態の情報記憶システム 2 では、図 9 のステップ S 1 1 3 の処理において、交代処理は、アクセス範囲内で正確な情報が得られなかったセクタ 1 1 1 a と、隣接範囲内で正確な情報が得られなかったセクタ 1 1 1 a との双方について行われる。

【 0 2 2 7 】

これにより、今回、記憶情報の読出しに失敗したセクタ 1 1 1 a だけでなく、そのセクタ 1 1 1 a に隣接して、記憶情報の読出しに失敗する可能性の高いセクタ 1 1 1 a についても交代処理が行われる。従って、情報記憶システム 2 におけるパフォーマンスの低下が、一層抑えられることとなっている。このことは、本件の情報記憶システムに対し、以下の応用形態が好適であることを意味している。

30

【 0 2 2 8 】

この応用形態では、上記交代部が、上記試行区画および上記隣接区画の双方について正確な情報が得られない場合のその試行区画およびその隣接区画について上記交代区画を割り当てるものとなっている。

【 0 2 2 9 】

本実施形態において、図 6 のステップ S 1 1 1 ~ S 1 1 3 の処理を実行しているリカバリリミット処理部 7 5 は、この応用形態における交代部の一例にも相当している。

40

【 0 2 3 0 】

尚、本実施形態では、上記のように、パフォーマンスの低下を一層抑えるために、アクセス範囲内で正確な情報が得られなかったセクタ 1 1 1 a と、隣接範囲内で正確な情報が得られなかったセクタ 1 1 1 a との双方について交替処理が行われる。しかしながら、本件の情報記憶システムは、この形態に限るものではない。本件の情報記憶システムは、例えば、アクセス範囲内で正確な情報が得られなかったセクタ 1 1 1 a のみについて交代処理が行われるものであっても良い。

【 0 2 3 1 】

また、上記では、本件の情報記憶システムの一例として、複数の HDD 1 0 0 を備えた RAID システム 2 を例示した。しかしながら、本件の情報記憶システムは、この形態に

50

限るものではない。本件の情報記憶システムは、HDD 100を1つだけ備えた情報記憶システムであっても良い。

【0232】

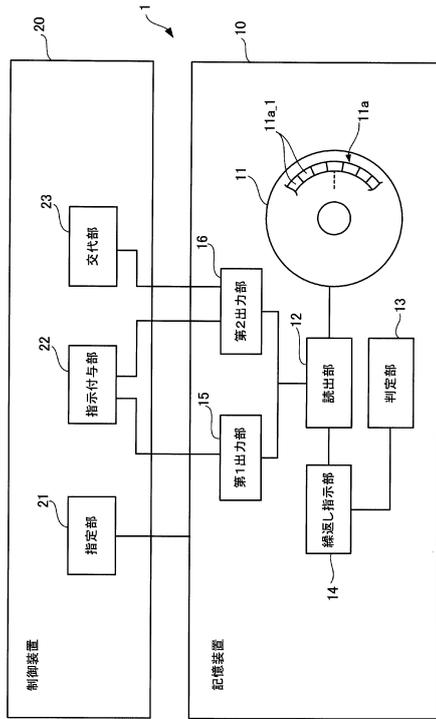
また、上記では、本件の情報記憶システムにおける記憶装置の一例としてHDD 100を例示した。しかしながら、本件の情報記憶システムにおける記憶装置はHDDに限るものではない。この記憶装置は、例えば、光磁気ディスクやDVD等のドライブ装置であっても良い。

【符号の説明】

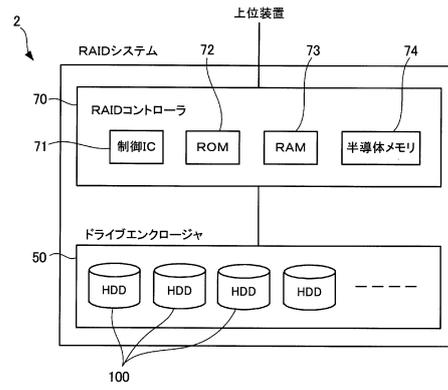
【0233】

1	情報記憶システム	10
2	RAIDシステム	
10	記憶装置	
11	記憶媒体	
11a, 111	トラック	
11a_1	記憶区画	
12	読出部	
13, 122	判定部	
14, 123	繰返し指示部	
15, 124	第1出力部	
16, 125	第2出力部	20
20	制御装置	
21	指定部	
22	指示付与部	
23	交代部	
50	ドライブエンクロージャ	
70	RAIDコントローラ	
71, 102	制御IC	
72, 103	ROM	
73, 105	RAM	
74, 106	半導体メモリ	30
75	リカバリリミット処理部	
76	媒体データ格納部	
77	データ正当性分析部	
78	アドレス保持部	
100	HDD	
101	ヘッド	
103	ECC処理回路	
111	磁気ディスク	
111a	セクタ	
121	アクセス部	40

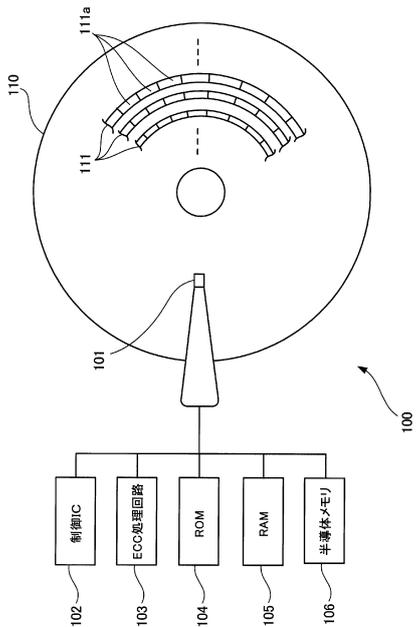
【図1】



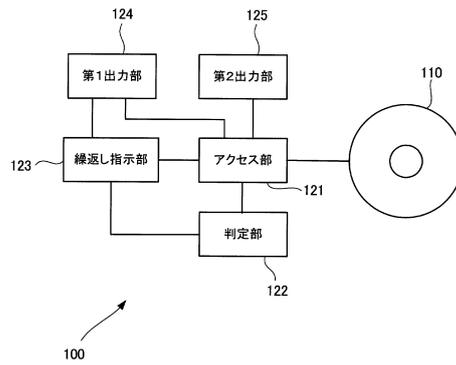
【図2】



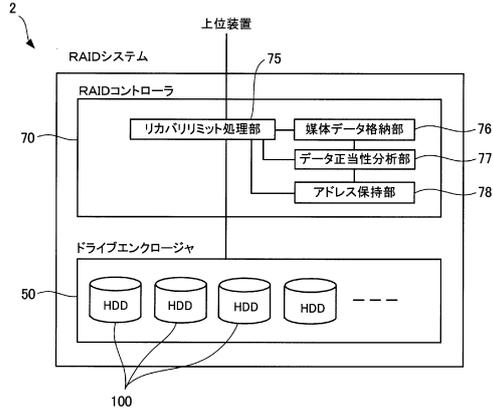
【図3】



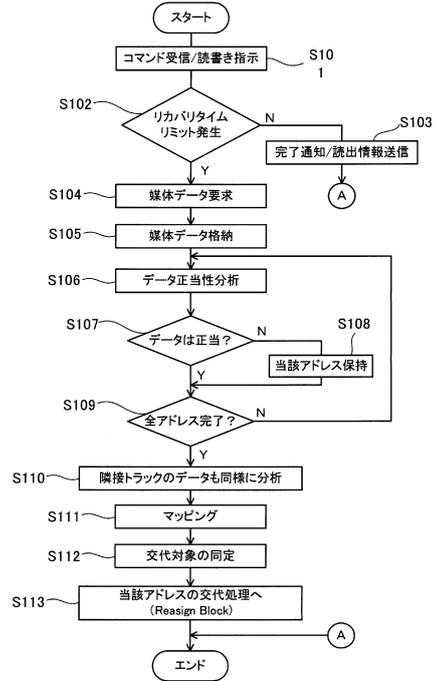
【図4】



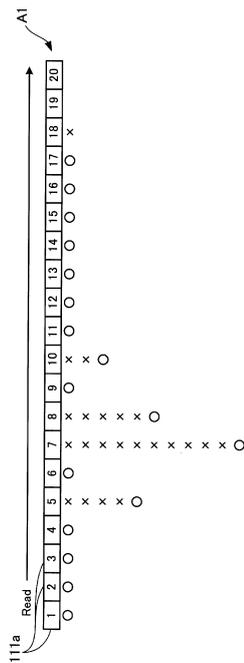
【図5】



【図6】



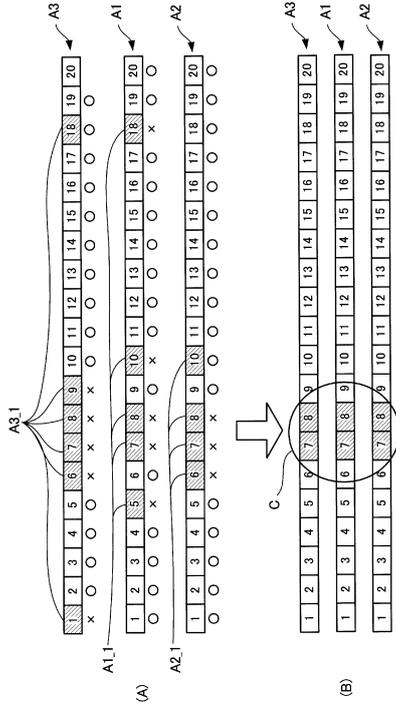
【図7】



【図8】



【 図 9 】



フロントページの続き

(51) Int.Cl.	F I		
	G 1 1 B	20/18	5 5 2 F
	G 1 1 B	20/18	5 7 2 B
	G 1 1 B	20/18	5 7 2 F
	G 1 1 B	20/12	
	G 1 1 B	20/10	C
	G 1 1 B	20/10	3 0 1 Z

- (56) 参考文献 特開2001-331377(JP, A)
特開2008-010113(JP, A)
特開2009-252001(JP, A)
特開2004-110946(JP, A)
特開平08-180612(JP, A)
特開2004-171671(JP, A)

(58) 調査した分野(Int.Cl., DB名)

G 1 1 B 2 0 / 1 8
G 0 6 F 3 / 0 6
G 1 1 B 2 0 / 1 0
G 1 1 B 2 0 / 1 2