(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第4156499号
(P4156499)
(45)発行日　平成20年9月24日(2008.9.24)　(24)登録日　平成20年7月18日(2008.7.18)

(51)Int.Cl.　F I
*G06F 3/06 (2006.01)*　G06F 3/06 304R
G06F 3/06 305C
G06F 3/06 540

請求項の数 31　(全 27 頁)

(21)出願番号　特願2003-400517(P2003-400517)
(22)出願日　平成15年11月28日(2003.11.28)
(65)公開番号　特開2005-165443(P2005-165443A)
(43)公開日　平成17年6月23日(2005.6.23)
審査請求日　平成18年11月16日(2006.11.16)

早期審査対象出願

前置審査

(73)特許権者　000005108
株式会社日立製作所
東京都千代田区丸の内一丁目６番６号
(74)代理人　100079108
弁理士　稲葉　良幸
(74)代理人　100093861
弁理士　大賀　眞司
(72)発明者　加納　東
神奈川県小田原市中里３２２番２号　株式会社日立製作所　ＲＡＩＤシステム事業部内
(72)発明者　小河　卓二
神奈川県小田原市中里３２２番２号　株式会社日立製作所　ＲＡＩＤシステム事業部内

最終頁に続く

(54)【発明の名称】ディスクアレイ装置

(57)【特許請求の範囲】
【請求項１】
第一の情報処理装置から書き込み要求又は／及び読み出し要求を受信して、第一のＦｉｂｒｅ　Ｃｈａｎｎｅｌ－Ａｒｂｉｔｒａｔｅｄ　Ｌｏｏｐ（ＦＣ－ＡＬ）を介してＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ　Ａｒｒａｙ　ｏｆ　Ｉｎｅｘｐｅｎｓｉｖｅ　Ｄｉｓｋｓ）方式に従ってデータの書き込み又は／及び読み出しを制御する第一のディスクコントローラと、
前記第一の情報処理装置又は／及び第二の情報処理装置から書き込み要求又は／及び読み出し要求を受信して、第二のＦＣ－ＡＬを介してＲＡＩＤ方式に従ってデータの書き込み又は／及び読み出しを制御する第二のディスクコントローラと、
前記第一のＦＣ－ＡＬの第一のサーキット及び前記第二のＦＣ－ＡＬの第二のサーキットに接続され且つ前記第一のディスクコントローラ及び前記第二のディスクコントローラにおける書き込み又は読み出し制御に基づいて送受信されるＦｉｂｒｅ　Ｃｈａｎｎｅｌ（ＦＣ）プロトコルに従った信号とシリアルＡＴＡ　Ａｔｔａｃｈｍｅｎｔ（シリアルＡＴＡ）プロトコルに従った信号との間の信号変換を行う変換器と、前記変換器によって変換済みのシリアルＡＴＡプロトコルに従った信号によって受信されるデータを記憶するシリアルＡＴＡディスクドライブと、を有するシリアルＡＴＡディスクドライブユニットと、
複数の前記シリアルＡＴＡディスクドライブユニットと、前記複数のシリアルＡＴＡディスクドライブユニットへの電力の供給を行う第一の電源モジュールと、を有する第一の筐体と

を備え、
前記第一のサーキットは、前記第一のＦＣ－ＡＬ上の経路をバイパスする機能を有し、
前記第二のサーキットは、前記第二のＦＣ－ＡＬ上の経路をバイパスする機能を有する
ことを特徴とするディスクアレイ装置。
【請求項２】
請求項１に記載のディスクアレイ装置において、
前記第一のサーキットは、ポートバイパスサーキットであり、
前記第一の筐体は、前記第一のＦＣ－ＡＬによる信号の通信に利用され且つ複数の前記第一のサーキットを有する第一の回路基板を有しており、
前記複数のシリアルＡＴＡディスクドライブユニットは、前記第一の筐体内において前記第一の回路基板に接続されるものである
ことを特徴とするディスクアレイ装置。
【請求項３】
請求項１又は２に記載のディスクアレイ装置において、
前記第二のサーキットは、ポートバイパスサーキットであり、
前記第一の筐体は、前記第二のＦＣ－ＡＬによる信号の通信に利用され且つ複数の前記第二のサーキットを有する第二の回路基板を有しており、
前記複数のシリアルＡＴＡディスクドライブユニットは、前記第一の筐体内において前記第二の回路基板に接続されるものである
ことを特徴とするディスクアレイ装置。
【請求項４】
請求項１乃至３のいずれかに記載のディスクアレイ装置において、
前記複数のシリアルＡＴＡディスクドライブユニットは、少なくとも一つのＲＡＩＤグループを形成しており、
前記ＲＡＩＤグループは、前記第一の情報処理装置又は／及び前記第二の情報処理装置からの書き込み要求又は読み出し要求の宛先となる複数の論理ボリューム識別子に対応付けられるものである
ことを特徴とするディスクアレイ装置。
【請求項５】
請求項１乃至３のいずれかに記載のディスクアレイ装置において、
ＦＣディスクドライブを有するＦＣディスクドライブユニットと、
複数の前記ＦＣディスクドライブユニットと、前記複数のＦＣディスクドライブユニットへの電力の供給を行う第二の電源モジュールと、を有する第二の筐体と、を有しており、
前記複数のシリアルＡＴＡディスクドライブは、少なくとも一つの第一のＲＡＩＤグループを形成しており、
前記複数のＦＣディスクドライブは、少なくとも一つの第二のＲＡＩＤグループを形成しているものである
ことを特徴とするディスクアレイ装置。
【請求項６】
請求項５に記載のディスクアレイ装置において、
前記第一のディスクコントローラは、前記第一のＲＡＩＤグループ内のシリアルＡＴＡディスクドライブと、前記第二のＲＡＩＤグループ内のＦＣディスクドライブと、のいずれのディスクドライブへのデータの書き込み又は／及び読み出しも可能である
ことを特徴とするディスクアレイ装置。
【請求項７】
請求項１乃至６のいずれかに記載のディスクアレイ装置において、
前記複数のシリアルＡＴＡディスクドライブユニットは、前記第一のＦＣ－ＡＬに接続される第一の接続部と、前記第二のＦＣ－ＡＬに接続される第二の接続部と、を有するものである

ことを特徴とするディスクアレイ装置。
【請求項８】
請求項５又は６のいずれかに記載のディスクアレイ装置において、
前記第一のディスクコントローラ又は／及び前記第二のディスクコントローラは、前記第二のＲＡＩＤグループ内のＦＣディスクドライブに格納されるデータには行わないデータの信頼性のチェックを、前記第一のＲＡＩＤグループ内のシリアルＡＴＡディスクドライブ内に格納されるデータへ行うものである
ことを特徴とするディスクアレイ装置。
【請求項９】
請求項５又は６のいずれかに記載のディスクアレイ装置において、
前記第一のディスクコントローラ又は／及び前記第二のディスクコントローラは、前記第一のＲＡＩＤグループ内のシリアルＡＴＡディスクドライブへのデータの書き込み又は読み出しの際に、前記第二のＲＡＩＤグループ内のＦＣディスクドライブへのデータの書き込み又は読み出し制御と異なる制御を行うものである
ことを特徴とするディスクアレイ装置。
【請求項１０】
第一の情報処理装置から書き込み要求を受信して、第一のＦｉｂｒｅ Ｃｈａｎｎｅｌ－Ａｒｂｉｔｒａｔｅｄ Ｌｏｏｐ（ＦＣ－ＡＬ）を介してＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）方式に従ってデータの書き込みを制御する第一のディスクコントローラと、
前記第一の情報処理装置又は／及び第二の情報処理装置から書き込み要求を受信して、第二のＦＣ－ＡＬを介してＲＡＩＤ方式に従ってデータの書き込みを制御する第二のディスクコントローラと、
前記第一のＦＣ－ＡＬの第一のサーキットに接続される第一の接続部と、前記第二のＦＣ－ＡＬの第二のサーキットに接続される第二の接続部と、前記第一のディスクコントローラから前記第一の接続部を介して受信されるＦｉｂｒｅ Ｃｈａｎｎｅｌ（ＦＣ）プロトコルに従った信号及び前記第二のディスクコントローラから前記第二の接続部を介して受信されるＦＣプロトコルに従った信号の両方をシリアルＡＴ Ａｔｔａｃｈｍｅｎｔ（シリアルＡＴＡ）プロトコルに従った信号へ変換を行う変換器と、前記変換器によって変換済みのシリアルＡＴＡプロトコルに従った信号によって受信されるデータを記憶するシリアルＡＴＡディスクドライブと、を有する第一のディスクドライブユニットと、
複数の前記第一のディスクドライブユニットと、を有する第一の筐体と
を備え、
前記第一のサーキットは、前記第一のＦＣ－ＡＬ上の経路をバイパスする機能を有し、
前記第二のサーキットは、前記第二のＦＣ－ＡＬ上の経路をバイパスする機能を有する
ことを特徴とするディスクアレイ装置。
【請求項１１】
第一の情報処理装置から書き込み要求を受信して、第一のＦｉｂｒｅ Ｃｈａｎｎｅｌ－Ａｒｂｉｔｒａｔｅｄ Ｌｏｏｐ（ＦＣ－ＡＬ）を介してＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）方式に従ってデータの書き込みを制御する第一のディスクコントローラと、
前記第一の情報処理装置又は／及び第二の情報処理装置から書き込み要求を受信して、第二のＦＣ－ＡＬを介してＲＡＩＤ方式に従ってデータの書き込みを制御する第二のディスクコントローラと、
前記第１のＦＣ-ＡＬの第一のサーキット及び前記第二のＦＣ-ＡＬの第二のサーキットに接続され前記第一のディスクコントローラから第一の接続部を介して受信されるＦｉｂｒｅ Ｃｈａｎｎｅｌ（ＦＣ）プロトコルに従った信号をシリアルＡＴ Ａｔｔａｃｈｍｅｎｔ（シリアルＡＴＡ）プロトコルに従った信号へ変換しかつ前記第二のディスクコントローラから第二の接続部を介して受信されるＦＣプロトコルに従った信号をシリアルＡＴＡプロトコルに従った信号へ変換する変換器と、前記変換器によって変換済みのシリア

ルＡＴＡプロトコルに従った信号によって受信されるデータを記憶するシリアルＡＴＡディスクドライブと、を有する第一のディスクドライブユニットと、
複数の前記第一のディスクドライブユニットを有する第一の筐体と、
を備え、
前記第一のサーキットは、前記第一のＦＣ－ＡＬ上の経路をバイパスする機能を有し、
前記第二のサーキットは、前記第二のＦＣ－ＡＬ上の経路をバイパスする機能を有する
ことを特徴とするディスクアレイ装置。
【請求項１２】
請求項１０又は１１に記載のディスクアレイ装置において、
前記第一の筐体は、前記複数の第一のディスクドライブユニットへの電力の供給を行う第一の電源モジュールを備える
ことを特徴とするディスクアレイ装置。
【請求項１３】
請求項１０乃至１２のいずれかに記載のディスクアレイ装置において、
前記第一のサーキットは、ポートバイパスサーキットであり、
前記第一の筐体は、前記第一のＦＣ-ＡＬによる信号の通信に利用され、複数の前記第一のサーキットを有する第一の回路基板を有しており、
前記複数の第一のディスクドライブユニットは、各々の前記第一の接続部を介して前記第一の回路基板に接続されるものである
ことを特徴とするディスクアレイ装置。
【請求項１４】
請求項１０乃至１３のいずれかに記載のディスクアレイ装置において、
前記第二のサーキットは、ポートバイパスサーキットであり、
前記第一の筐体は、前記第二のＦＣ－ＡＬによる信号の通信に利用され、複数の前記第二のサーキットを有する第二の回路基板を有しており、
前記複数の第一のディスクドライブユニットは、各々の前記第二の接続部を介して前記第二の回路基板に接続されるものである
ことを特徴とするディスクアレイ装置。
【請求項１５】
請求項１０乃至１４のいずれかに記載のディスクアレイ装置において、
前記複数の第一のディスクドライブユニットは、少なくとも一つのＲＡＩＤグループを形成しており、
前記ＲＡＩＤグループは、前記第一の情報処理装置又は／及び前記第二の情報処理装置からの書き込み要求の宛先となる複数の論理ボリューム識別子に対応付けられるものである
ことを特徴とするディスクアレイ装置。
【請求項１６】
請求項１０乃至１４のいずれかに記載のディスクアレイ装置において、
ＦＣディスクドライブを有する第二のディスクドライブユニットと、
複数の前記第二のディスクドライブユニットを有する第二の筐体と、を有しており、
前記複数のシリアルＡＴＡディスクドライブは、少なくとも一つの第一のＲＡＩＤグループを形成しており、
前記複数のＦＣディスクドライブは、少なくとも一つの第二のＲＡＩＤグループを形成しているものである
ことを特徴とするディスクアレイ装置。
【請求項１７】
請求項１６に記載のディスクアレイ装置において、
前記第一のディスクコントローラは、前記第一のＲＡＩＤグループ内のシリアルＡＴＡディスクドライブと、前記第二のＲＡＩＤグループ内のＦＣディスクドライブと、のいずれのディスクドライブへのデータの書き込みも可能である

ことを特徴とするディスクアレイ装置。
【請求項１８】
請求項１０乃至１７のいずれかに記載のディスクアレイ装置において、
前記第一の筐体は、さらに、
前記第一のディスクコントローラに前記第一のＦＣ－ＡＬを介して接続される第一のコネクタと、
前記第二のディスクコントローラに前記第二のＦＣ－ＡＬを介して接続される第一のコネクタと、を有する
ことを特徴とするディスクアレイ装置。
【請求項１９】
第一の情報処理装置から受信される書き込み要求及び読み出し要求に応じて、第一のＦｉｂｒｅ Ｃｈａｎｎｅｌ－Ａｒｂｉｔｒａｔｅｄ Ｌｏｏｐ（ＦＣ－ＡＬ）を介してＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）方式に従ってデータの書き込み及び読み出しを制御する第一のディスクコントローラと、
前記第一の情報処理装置及び第二の情報処理装置の少なくとも１つから受信される書き込み要求及び読み出し要求に応じて、第二のＦＣ－ＡＬを介してＲＡＩＤ方式に従ってデータの書き込み及び読み出しを制御する第二のディスクコントローラと、
複数のディスクドライブユニットを含む増設筐体と、
を備え、
前記複数のディスクドライブユニットの各々は、変換器とシリアルＡＴ Ａｔｔａｃｈｍｅｎｔ（シリアルＡＴＡ）ディスクドライブとを備えており、
前記変換器は、前記第一のＦＣ－ＡＬの第一のサーキットと前記第二のＦＣ－ＡＬの第二のサーキットとに接続され、前記第一のＦＣ－ＡＬ及び前記第二のＦＣ－ＡＬのファイバチャネルプロトコルと前記シリアルＡＴＡディスクドライブのシリアルＡＴＡプロトコルとの間のプロトコルの変換を行うものであり、
複数のディスクドライブユニットの各々は、前記第一のディスクコントローラと前記第二のディスクコントローラとの制御に応じてデータが書込まれ又は/及び読み出されるものであり、
前記第一のサーキットは、前記第一のＦＣ－ＡＬ上の経路をバイパスする機能を有し、
前記第二のサーキットは、前記第二のＦＣ－ＡＬ上の経路をバイパスする機能を有する
ことを特徴とするディスクアレイシステム。
【請求項２０】
請求項１９に記載のディスクアレイシステムにおいて、
前記増設筐体は複数の電力供給モジュールを含み、前記複数の電力供給モジュールの各々は、前記複数のディスクドライブユニットに電力を供給するように制御するものであり、
前記第一のディスクコントローラ及び前記第二のディスクコントローラは、前記増設筐体の外に備えられるものである、
ことを特徴とするディスクアレイシステム。
【請求項２１】
請求項２０に記載のディスクアレイシステムにおいて、
前記増設筐体は、第一のコネクタと、第二のコネクタとを備えており、
前記第一のコネクタは、前記第一のディスクコントローラと前記複数のディスクドライブユニットとの間のデータの中継に利用される前記第一のＦＣ－ＡＬに接続されるものであり、
前記第二のコネクタは、前記第二のディスクコントローラと前記複数のディスクドライブユニットとの間のデータの中継に利用される前記第二のＦＣ－ＡＬに接続されるものである、
ことを特徴とするディスクアレイシステム。

【請求項２２】
請求項１９又は２０に記載のディスクアレイシステムにおいて、
前記ディスクアレイシステムは、前記第一のディスクコントローラと前記第二のディスクコントローラとを含む基本筐体を備えており、
前記第一及び第二のサーキットは、ポートバイパスサーキットであり、
前記増設筐体は、複数の前記第一及び第二のサーキットを備えており、
前記複数の前記第一及び第二のサーキットのうちの一つは、前記複数のディスクドライブユニットのうち少なくとも一つのディスクドライブユニットに対応する前記変換器に接続されるものである、
ことを特徴とするディスクアレイシステム。
【請求項２３】
請求項１９乃至２２のいずれかに記載のディスクアレイシステムにおいて、
前記第一及び第二のサーキットは、ポートバイパスサーキットであり、
前記増設筐体は、前記第一のディスクコントローラに接続される前記第一のサーキットと、前記第二のディスクコントローラに接続される前記第二のサーキットと、を備えている、
ことを特徴とするディスクアレイシステム。
【請求項２４】
請求項１９乃至２２のいずれかに記載のディスクアレイシステムにおいて、
前記第一及び第二のサーキットは、ポートバイパスサーキットであり、
前記増設筐体は、回路基板を備えており、
前記回路基板は、複数の前記第一及び第二のサーキットを備えており、前記複数のディスクドライブユニットの各々に対応する前記変換器に接続されるものである、
ことを特徴とするディスクアレイシステム。
【請求項２５】
請求項１９乃至２４のいずれかに記載のディスクアレイシステムにおいて、
前記第一のディスクコントローラと前記第二のディスクコントローラとを有する基本筐体を備える、
ことを特徴とするディスクアレイシステム。
【請求項２６】
請求項１９乃至２５のいずれかに記載のディスクアレイシステムにおいて、
前記第一のディスクコントローラによって実行され、前記複数のディスクドライブユニットのうちの２つ以上のディスクドライブユニットを用いて第一のＲＡＩＤグループを形成するように制御する第一のプログラムと、
前記第二のディスクコントローラによって実行され、前記複数のディスクドライブユニットのうちの２つ以上のディスクドライブユニットを用いて第二のＲＡＩＤグループを形成するように制御する第二のプログラムとを備える、
ことを特徴とするディスクアレイシステム。
【請求項２７】
請求項１９乃至２６のいずれかに記載のディスクアレイシステムにおいて、
前記ディスクアレイシステムは、ファイバチャネルインタフェースを有するファイバチャネルディスクドライブを各々に有する複数の他のディスクドライブユニットと、前記複数の他のディスクドライブユニットに対する電力の供給を制御する複数の第二の電力供給モジュールと、を備える他の増設筐体とを備える、
ことを特徴とするディスクアレイシステム。
【請求項２８】
請求項２７に記載のディスクアレイシステムにおいて、
前記第一のディスクコントローラは、前記複数のディスクドライブユニット内の複数の前記シリアルＡＴＡディスクドライブ及び前記複数の他のディスクドライブユニット内の複数の前記ファイバチャネルディスクドライブに対して、データの読み出し及び書き込み

を制御するものである、
ことを特徴とするディスクアレイシステム。
【請求項２９】
請求項２８に記載のディスクアレイシステムにおいて、
前記第二のディスクコントローラは、前記複数のディスクドライブユニット内の複数の前記シリアルＡＴＡディスクドライブ及び前記複数の他のディスクドライブユニット内の複数の前記ファイバチャネルディスクドライブに対して、データの読み出し及び書き込みを制御することが可能なものである、
ことを特徴とするディスクアレイシステム。
【請求項３０】
請求項１９乃至２９のいずれかに記載のディスクアレイシステムにおいて、
前記第一のディスクコントローラによって実行され、第一の論理ユニット番号（ＬＵＮ）に対応する第一の論理ボリュームを形成するように制御する第一のコンピュータプログラムと、
前記第二のディスクコントローラによって実行され、第二のＬＵＮに対応する第二の論理ボリュームを形成するように制御する第二のコンピュータプログラムと、
を備えることを特徴とするディスクアレイシステム。
【請求項３１】
請求項１９乃至３０いずれかに記載のディスクアレイシステムにおいて、
前記複数のディスクドライブユニットの各々は、前記増設筐体に対して挿入されるものであり、かつ前記増設筐体から取り外すことが可能なものである、
ことを特徴とするディスクアレイシステム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、ディスクアレイ装置及びディスクアレイ装置の制御方法に関する。
【背景技術】
【０００２】
近年、ディスクアレイ装置における記憶容量の増大に伴い、情報処理システムにおける重要性は益々高まってきている。そこで、情報処理装置等からのデータ入出力要求に対して、要求された位置に正しくデータの書き込みを行うこと、読み出したデータが不正である場合にはそれを検知することが重要である。
【０００３】
特許文献１においては、磁気ディスク装置に２つのヘッドを持たせ、同一のデータを２つのヘッドから読み出して比較することにより、磁気ディスク装置における書き込み及び読み出しの信頼性を高める方法が開示されている。
【特許文献１】特開平５－１５０９０９号公報
【発明の開示】
【発明が解決しようとする課題】
【０００４】
特許文献１の方法をディスクアレイ装置に適用する場合、各磁気ディスクにヘッドを２つ持たせる必要があるため、ハードディスクドライブの製造単価が高くなってしまう。そこで、ヘッドの追加等の物理的な構造の変更を行うことなく、ハードディスクドライブにおける信頼性を高める方法が求められている。
【０００５】
また、ディスクアレイ装置においては、ファイバチャネルのハードディスクドライブに加えて、シリアルＡＴＡやパラレルＡＴＡ等のハードディスクドライブ等も利用されはじめている。これは、シリアルＡＴＡやパラレルＡＴＡ等のハードディスクドライブは、ファイバチャネルのハードディスクドライブと比較して信頼性は劣るが価格が低いためである。そこで、このようにファイバチャネルとシリアルＡＴＡ等の規格のハードディスクド

ライブを組み合わせて構成するディスクアレイ装置において、ファイバチャネル以外のハードディスクドライブにおける信頼性を高める方法が求められている。

【０００６】

本発明は上記課題を鑑みてなされたものであり、ディスクアレイ装置及びディスクアレイ装置の制御方法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】

上記目的を達成するため本発明のうち主たる発明に係るディスクアレイ装置は、第一の情報処理装置から書き込み要求又は／及び読み出し要求を受信して、第一のＦｉｂｒｅ Ｃｈａｎｎｅｌ－Ａｒｂｉｔｒａｔｅｄ Ｌｏｏｐ（ＦＣ－ＡＬ）を介してＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）方式に従ってデータの書き込み又は／及び読み出しを制御する第一のディスクコントローラと、前記第一の情報処理装置又は／及び第二の情報処理装置から書き込み要求又は／及び読み出し要求を受信して、第二のＦＣ－ＡＬを介してＲＡＩＤ方式に従ってデータの書き込み又は／及び読み出しを制御する第二のディスクコントローラと、前記第一のＦＣ－ＡＬの第一のサーキット及び前記第二のＦＣ－ＡＬの第二のサーキットに接続され前記第一のディスクコントローラ及び前記第二のディスクコントローラにおける書き込み又は読み出し制御に基づいて送受信されるＦｉｂｒｅ Ｃｈａｎｎｅｌ（ＦＣ）プロトコルに従った信号とシリアルＡＴ Ａｔｔａｃｈｍｅｎｔ（シリアルＡＴＡ）プロトコルに従った信号との間の信号変換を行う変換器と、前記変換器によって変換済みのシリアルＡＴＡプロトコルに従った信号によって受信されるデータを記憶するシリアルＡＴＡディスクドライブと、を有するシリアルＡＴＡディスクドライブユニットと、複数の前記シリアルＡＴＡディスクドライブユニットと、前記複数のシリアルＡＴＡディスクドライブユニットへの電力の供給を行う第一の電源モジュールと、を有する第一の筐体とを備え、前記第一のサーキットは、前記第一のＦＣ－ＡＬ上の経路をバイパスする機能を有し、前記第二のサーキットは、前記第二のＦＣ－ＡＬ上の経路をバイパスする機能を有することを特徴とする。

【０００８】

また本発明においては、第一の情報処理装置から書き込み要求を受信して、第一のＦｉｂｒｅ Ｃｈａｎｎｅｌ－Ａｒｂｉｔｒａｔｅｄ Ｌｏｏｐ（ＦＣ－ＡＬ）を介してＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）方式に従ってデータの書き込みを制御する第一のディスクコントローラと、前記第一の情報処理装置又は／及び第二の情報処理装置から書き込み要求を受信して、第二のＦＣ－ＡＬを介してＲＡＩＤ方式に従ってデータの書き込みを制御する第二のディスクコントローラと、前記第一のＦＣ－ＡＬの第一のサーキットに接続される第一の接続部と、前記第二のＦＣ－ＡＬの第二のサーキットに接続される第二の接続部と、前記第一のディスクコントローラから前記第一の接続部を介して受信されるＦｉｂｒｅ Ｃｈａｎｎｅｌ（ＦＣ）プロトコルに従った信号及び前記第二のディスクコントローラから前記第二の接続部を介して受信されるＦＣプロトコルに従った信号の両方をシリアルＡＴ Ａｔｔａｃｈｍｅｎｔ（シリアルＡＴＡ）プロトコルに従った信号へ変換を行う変換器と、前記変換器によって変換済みのシリアルＡＴＡプロトコルに従った信号によって受信されるデータを記憶するシリアルＡＴＡディスクドライブと、を有する第一のディスクドライブユニットと、複数の前記第一のディスクドライブユニットと、を有する第一の筐体とを備え、前記第一のサーキットは、前記第一のＦＣ－ＡＬ上の経路をバイパスする機能を有し、前記第二のサーキットは、前記第二のＦＣ－ＡＬ上の経路をバイパスする機能を有することを特徴とする。

【０００９】

さらに本発明においては、第一の情報処理装置から書き込み要求を受信して、第一のＦｉｂｒｅ Ｃｈａｎｎｅｌ－Ａｒｂｉｔｒａｔｅｄ Ｌｏｏｐ（ＦＣ－ＡＬ）を介してＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）方式に従ってデータの書き込みを制御する第一のディスクコントローラと、前記第一の

情報処理装置又は／及び第二の情報処理装置から書き込み要求を受信して、第二のＦＣ－ＡＬを介してＲＡＩＤ方式に従ってデータの書き込みを制御する第二のディスクコントローラと、前記第１のＦＣ-ＡＬの第一のサーキット及び前記第二のＦＣ-ＡＬの第二のサーキットに接続され前記第一のディスクコントローラから第一の接続部を介して受信されるＦｉｂｒｅ Ｃｈａｎｎｅｌ（ＦＣ）プロトコルに従った信号をシリアルＡＴ Ａｔｔａｃｈｍｅｎｔ（シリアルＡＴＡ）プロトコルに従った信号へ変換し、かつ前記第二のディスクコントローラから第二の接続部を介して受信されるＦＣプロトコルに従った信号をシリアルＡＴＡプロトコルに従った信号へ変換する変換器と、前記変換器によって変換済みのシリアルＡＴＡプロトコルに従った信号によって受信されるデータを記憶するシリアルＡＴＡディスクドライブと、を有する第一のディスクドライブユニットと、複数の前記第一のディスクドライブユニットを有する第一の筐体と、を備え、前記第一のサーキットは、前記第一のＦＣ－ＡＬ上の経路をバイパスする機能を有し、前記第二のサーキットは、前記第二のＦＣ－ＡＬ上の経路をバイパスする機能を有することを特徴とする。

【００１０】

さらに本発明においては、第一の情報処理装置から受信される書き込み要求及び読み出し要求に応じて、第一のＦｉｂｒｅ Ｃｈａｎｎｅｌ－Ａｒｂｉｔｒａｔｅｄ Ｌｏｏｐ（ＦＣ－ＡＬ）を介してＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）方式に従ってデータの書き込み及び読み出しを制御する第一のディスクコントローラと、前記第一の情報処理装置及び第二の情報処理装置の少なくとも１つから受信される書き込み要求及び読み出し要求に応じて、第二のＦＣ－ＡＬを介してＲＡＩＤ方式に従ってデータの書き込み及び読み出しを制御する第二のディスクコントローラと、複数のディスクドライブユニットを含む増設筐体と、を備え、前記複数のディスクドライブユニットの各々は、変換器とシリアルＡＴ Ａｔｔａｃｈｍｅｎｔ（シリアルＡＴＡ）ディスクドライブとを備えており、前記変換器は、前記第一のＦＣ－ＡＬの第一のサーキットと前記第二のＦＣ－ＡＬの第二のサーキットとに接続され、前記第一のＦＣ－ＡＬ及び前記第二のＦＣ－ＡＬのファイバチャネルプロトコルと前記シリアルＡＴＡディスクドライブのシリアルＡＴＡプロトコルとの間のプロトコルの変換を行うものであり、複数のディスクドライブユニットの各々は、前記第一のディスクコントローラと前記第二のディスクコントローラとの制御に応じてデータが書込まれ又は/及び読み出されるものであり、前記第一のサーキットは、前記第一のＦＣ－ＡＬ上の経路をバイパスする機能を有し、前記第二のサーキットは、前記第二のＦＣ－ＡＬ上の経路をバイパスする機能を有することを特徴とする。

【００１１】

その他、本願が開示する課題、及びその解決方法は、発明を実施するための最良の形態の欄、及び図面により明らかにされる。

【発明の効果】

【００１２】

ディスクアレイ装置及びディスクアレイ装置の制御方法を提供することができる。

【発明を実施するための最良の形態】

【００１３】

＝＝装置構成＝＝

図１（ａ）は本発明の一実施例として説明するディスクアレイ装置１０の正面図であり、図１（ｂ）はディスクアレイ装置１０の背面図である。図２（ａ）は、ディスクアレイ装置１０に装着される基本筐体２０を正面側から見た斜視図であり、図２（ｂ）は基本筐体２０を背面側から見た斜視図である。図３（ａ）は、ディスクアレイ装置１０に装着される増設筐体３０を正面側から見た斜視図であり、図３（ｂ）は増設筐体３０を背面側から見た斜視図である。

【００１４】

図１（ａ），（ｂ）に示すように、ディスクアレイ装置１０は、ラックフレーム１１をベースとして構成される。ラックフレーム１１の内側左右側面の上下方向には、複数段に

わたって前後方向にマウントフレーム１２が形成され、このマウントフレーム１２に沿って基本筐体２０および増設筐体３０が引き出し式に装着される。図２（ａ），（ｂ）に示すように、基本筐体２０および増設筐体３０には、ディスクアレイ装置１０の各種機能を提供するボードやユニットが装着されている。

【００１５】

図２（ａ）に示すように、基本筐体２０の正面上段側には、ハードディスクドライブ５１が装填された複数のディスクドライブユニット５２が並べて装着されている。ハードディスクドライブ５１は、例えば、ＦＣ－ＡＬ規格、ＳＣＳＩ１（Small Computer System Interface 1）規格、ＳＣＳＩ２規格、ＳＣＳＩ３規格、パラレルＡＴＡ（AT Attachment）規格、シリアルＡＴＡ規格などの通信機能を提供する通信インタフェースを有するハードディスクドライブである。

【００１６】

また、基本筐体２０の正面下段側には、バッテリーユニット５３、ハードディスクドライブ５１の稼働状態などが表示される表示パネル５４、フレキシブルディスクドライブ５５が装着されている。バッテリーユニット５３には二次電池が内蔵されている。バッテリーユニット５３は、停電などによりＡＣ／ＤＣ電源５７からの電力供給が途絶えた場合に、ボードやユニットに電力を供給するバックアップ電源として機能する。表示パネル５４には、ハードディスクドライブ５１の稼働状態などを表示するＬＥＤランプなどの表示デバイスが設けられている。フレキシブルディスクドライブ５５は、メンテナンス用プログラムをロードする場合などに用いられる。

【００１７】

図２（ｂ）に示すように、基本筐体２０の背面上段側の両側面側には、１枚ずつ電源コントローラボード５６が装着されている。電源コントローラボード５６は、複数のハードディスクドライブ５１と通信可能に接続されている。電源コントローラボード５６と複数のハードディスクドライブ５１は、ループ状の通信経路、例えば、ＦＣ－ＡＬの方式（トポロジー）で通信を行う通信経路によって通信可能に接続されている。

【００１８】

電源コントローラボード５６は、ＡＣ／ＤＣ電源５７の状態監視やハードディスクドライブ５１の状態監視、ハードディスクドライブ５１の電源供給の制御、冷却装置の冷却能力の制御、表示パネル５４上の表示デバイスの制御、筐体各部の温度監視などを行う回路が実装されている。なお、冷却装置は、ディスクアレイ装置１０内や筐体２０，３０内を冷却する装置であり、例えば、インタークーラー、ヒートシンク、空冷式の冷却ファンなどである。電源コントローラボード５６にはファイバチャネルケーブルのコネクタ６７が設けられ、このコネクタにはファイバチャネルケーブル９１が接続される。

【００１９】

図２（ｂ）に示すように、基本筐体２０の背面上段側の前記２枚の電源コントローラボード５６に挟まれた空間には、ＡＣ／ＤＣ電源５７が２台並べて装着されている。ＡＣ／ＤＣ電源５７は、ハードディスクドライブ５１、ボード、ユニットなどに電源を供給する。ＡＣ／ＤＣ電源５７は、電源コントローラボード５６と接続されており、電源コントローラボード５６からの信号により各ハードディスクドライブ５１に電源を供給できるように設定されている。

【００２０】

なお、本実施の形態においては、各筐体２０，３０の電源供給に関するセキュリティを確保するために、基本筐体２０および増設筐体３０に電源コントローラボード５６とＡＣ／ＤＣ電源５７とを各２台ずつ冗長に装着させることとしているが、電源コントローラボード５６とＡＣ／ＤＣ電源５７とを各１台ずつ装着させることとしてもよい。

ＡＣ／ＤＣ電源５７には、ＡＣ／ＤＣ電源５７の出力をオン・オフするためのブレーカスイッチ６４が設けられている。

【００２１】

図２（ｂ）に示すように、ＡＣ／ＤＣ電源５７の下方には、２台の空冷式の冷却ファン

ユニット５８が並べて装着されている。冷却ファンユニット５８には、１台以上の冷却ファン６６が実装されている。冷却ファン６６は、筐体内に空気を流入・流出させることでハードディスクドライブ５１やＡＣ／ＤＣ電源５７などから発生する熱を筐体外部に排出する。なお、基本筐体２０や増設筐体３０、およびこれらに装着されているボードやユニットには、筐体２０，３０内に空気を循環させる通気路や通気口が形成され、冷却ファン６６により筐体２０内の熱が外部に効率よく排出される仕組みになっている。冷却ファン６６は、ハードディスクドライブ５１ごとに設けることとしてもよいが、チップやユニットの数を削減できることから筐体ごとに大きな冷却ファン６６を設けることが好ましい。

【００２２】

冷却ファンユニット５８は、コントローラボード５９もしくは電源コントローラボード５６と制御ライン４８で接続されており、冷却ファンユニット５８の冷却ファン６６の回転数は、この制御ライン４８を通じてコントローラボード５９もしくは電源コントローラボード５６により制御される。

【００２３】

図２（ｂ）に示すように、基本筐体２０の背面下段側には、１枚のコントローラボード５９が装着されている。コントローラボード５９には、基本筐体２０および増設筐体３０に装着されているハードディスクドライブ５１との間の通信インタフェースと、ハードディスクドライブ５１の動作の制御（例えば、ＲＡＩＤ方式による制御）やハードディスクドライブ５１の状態監視を行う回路などが実装されている。

【００２４】

なお、本実施の形態において、電源コントローラボード５６がハードディスクドライブ５１の電源供給の制御や冷却装置の冷却能力の制御を行うこととしているが、これらの制御をコントローラボード５９が行うこととしてもよい。

【００２５】

また、本実施例においては、コントローラボード５９は、情報処理装置３００との間の通信インタフェースの機能、例えば、ＳＣＳＩ規格やファイバチャネル規格の通信機能を提供する通信インタフェースボード６１や、ハードディスクドライブ５１への書き込みデータや読み出しデータが記憶されるキャッシュメモリ６２などを実装する形態としているが、これらを別のボートが実装する形態としてもよい。

【００２６】

コントローラボード５９に装着される通信インタフェースボード６１には、情報処理装置３００と接続するための、ファイバチャネル、Ethernet （登録商標）などのプロトコルで構築されたＳＡＮ（Storage Area Network）、ＬＡＮ（Local Area Network）、もしくは、ＳＣＳＩなどの所定のインタフェース規格に準拠した外部コネクタ６３が設けられ、ディスクアレイ装置１０は、このコネクタ６３に接続される通信ケーブル９２を介して情報処理装置３００と接続される。

【００２７】

なお、基本筐体２０のハードディスクドライブ５１の制御に関するセキュリティを確保するために、２枚のコントローラボード５９を冗長に装着させることとしてもよい。

【００２８】

図３（ａ）に示すように、増設筐体３０の正面側には、ハードディスクドライブ５１が収容された複数のディスクドライブユニット５２が並べて装着されている。図３（ｂ）に示すように、増設筐体３０の背面両側面側には、それぞれ一枚ずつ電源コントローラボード５６が装着されている。また、２枚の電源コントローラボード５６に挟まれた空間には、ＡＣ／ＤＣ電源５７が２台並べて装着されている。また、ＡＣ／ＤＣ電源５７の下方には、２台の冷却ファンユニット５８が並べて装着されている。ＡＣ／ＤＣ電源５７には、ＡＣ／ＤＣ電源５７の出力をオン・オフするためのブレーカスイッチ６４が設けられている。

【００２９】

本実施の形態においては、上述したように増設筐体３０の電源供給に関するセキュリテ

ィを確保するために、増設筐体３０に電源コントローラボード５６とＡＣ／ＤＣ電源５７とを各２台ずつ冗長に装着させることとしているが、電源コントローラボード５６とＡＣ／ＤＣ電源５７とを各１台ずつ装着させることとしてもよい。なお、ハードディスクドライブ５１の電源供給の制御や冷却装置の冷却能力の制御などの電源コントローラボード５６の機能をコントローラボード５９に実装することとしてもよい。

【００３０】

図４にディスクドライブユニット５２に収容されているハードディスクドライブ５１の構成の一例を示す。ハードディスクドライブ５１は、その筐体７０内に、磁気ディスク７３、アクチュエータ７１、スピンドルモータ７２、データの読み書きを行うヘッド７４、ヘッド７４等の機構部分を制御する機構制御回路７５、磁気ディスク７３へのデータの読み書き信号を制御する信号処理回路７６、通信インタフェース回路７７、各種コマンドやデータが入出力されるインタフェースコネクタ７９、電源コネクタ８０等を備えて構成される。なお、通信インタフェース回路７７には、データを一時的に格納するためのキャッシュメモリも含まれている。なお、後述するコントローラ５００におけるキャッシュメモリ６２と区別するため、ハードディスクドライブ５１が備えるキャッシュメモリをディスクキャッシュと称する。

【００３１】

ハードディスクドライブ５１は、例えば、コンタクトスタートストップ（ＣＳＳ：Contact Start Stop）方式の３．５インチサイズの磁気ディスクや、ロード／アンロード方式の２．５インチサイズの磁気ディスクなどを備える記憶装置である。３．５インチサイズの磁気ディスクは、例えば、ＳＣＳＩ１、ＳＣＳＩ２、ＳＣＳＩ３、ＦＣ－ＡＬなどの通信インタフェースを有している。一方、２．５インチサイズの磁気ディスクは、例えば、パラレルＡＴＡ、シリアルＡＴＡなどの通信インタフェースを有している。

【００３２】

２．５インチサイズの磁気ディスクをディスクアレイ装置１０の筐体２０，３０に収容する場合には、３．５インチの形状をした容器に収めるようにしてもよい。これにより、磁気ディスクの衝撃耐力性能を向上させることが可能となる。なお、２．５インチサイズの磁気ディスクと３．５インチサイズの磁気ディスクとは、通信インタフェースが異なるだけではなく、Ｉ／Ｏ性能、消費電力、寿命の点などで異なっている。２．５インチサイズの磁気ディスクは、３．５インチサイズの磁気ディスクに比べ、Ｉ／Ｏ性能が優れておらず、寿命が短い。しかし、３．５インチサイズの磁気ディスクに比べ、消費電力が少ないという点で優れている。

【００３３】

＝＝ディスクアレイ装置のハードウェア構成＝＝

図５は、本発明の一実施例として説明するディスクアレイ装置１０のハードウェア構成を示すブロック図である。

【００３４】

図５に示すように、ディスクアレイ装置１０には、ＳＡＮを介して情報処理装置３００が接続されている。情報処理装置３００は、例えば、パーソナルコンピュータ、ワークステーション、メインフレームコンピュータなどである。

【００３５】

ディスクアレイ装置１０は、前述したように、基本筐体２０と１つ又は複数の増設筐体３０を備えている。本実施の形態においては、基本筐体２０は、コントローラ５００、ハードディスクドライブ５１などを備えている。コントローラ５００は、チャネル制御部５０１、ディスク制御部５０２、ＣＰＵ５０３、メモリ５０４、キャッシュメモリ６２、及びデータコントローラ５０５などを備えており、前述のコントローラボード５９に実装されている。また、増設筐体３０は、ハードディスクドライブ５１などを備えている。基本筐体及び増設筐体のハードディスクドライブ５１は、ＦＣ－ＡＬ５０６により、ディスク制御部５０２と通信可能に接続されている。なお、ディスク制御部５０２とハードディスクドライブ５１との接続形態の詳細については後述する。

【００３６】
チャネル制御部５０１は情報処理装置３００との間で通信を行うインタフェースである。チャネル制御部５０１は、ファイバチャネルプロトコルに従ってブロックアクセス要求を受け付ける機能を有する。
【００３７】
ディスク制御部５０２は、ＣＰＵ５０３からの指示により、ハードディスクドライブ５１との間でデータのやりとりを行うインタフェースである。ディスク制御部５０２は、ハードディスクドライブ５１を制御するコマンドなどを規定するプロトコルに従ってハードディスクドライブ５１に対するデータ入出力要求を送信する機能を備える。
【００３８】
ＣＰＵ５０３は、ディスクアレイ装置１０の全体の制御を司るもので、メモリ５０４に格納されたマイクロプログラムを実行することにより、チャネル制御部５０１、ディスク制御部５０２、及びデータコントローラ５０６等の制御を行う。マイクロプログラムとは、図６に示すデータＲＥＡＤ処理６０１やデータＷＲＩＴＥ処理６０２などである。
【００３９】
キャッシュメモリ６２は、チャネル制御部５０１とディスク制御部５０２との間で授受されるデータを一時的に記憶するために用いられる。
【００４０】
データコントローラ５０５は、ＣＰＵ５０３の制御によりチャネル制御部５０１とキャッシュメモリ６２との間又はキャッシュメモリ６２とディスク制御部５０２との間のデータ転送を行うものである。
【００４１】
コントローラ５００は、ハードディスクドライブ５１をいわゆるＲＡＩＤ（Redundant Array of Inexpensive Disks）方式に規定されるＲＡＩＤレベル（例えば、０，１，５）で制御する機能を備えている。ＲＡＩＤ方式においては、複数のハードディスクドライブ５１が１つのグループ（以後、ＲＡＩＤグループと称する）として管理されている。ＲＡＩＤグループ上には、情報処理装置３００からのアクセス単位である論理ボリュームが形成されており、各論理ボリュームにはＬＵＮ（Logical Unit Number）と呼ばれる識別子が付与されている。ＲＡＩＤの構成情報は、図６に示すようにメモリ５０４にＲＡＩＤ構成テーブル６０３として記憶されており、データＲＥＡＤ処理６０１やデータＷＲＩＴＥ処理６０２の実行時にＣＰＵ５０３により参照される。
【００４２】
なお、ディスクアレイ装置は、以上に説明した構成のもの以外にも、例えば、ＮＦＳ（Network File System）などのプロトコルにより情報処理装置３００からファイル名指定によるデータ入出力要求を受け付けるように構成されたＮＡＳ（Network Attached Storage）として機能するものなどであってもよい。
【００４３】
＝＝ハードディスクドライブの接続形態＝＝
次に、コントローラ５００とハードディスクドライブ５１との接続形態について説明する。
図７は、基本筐体２０にファイバチャネルのハードディスクドライブ５１が収容されている場合における、ディスク制御部５０２と各ハードディスクドライブ５１との接続形態を示している。
【００４４】
ディスク制御部５０２は、ＦＣ－ＡＬ５０６で複数のハードディスクドライブ５１と接続されている。ＦＣ－ＡＬ５０６は、複数のＰＢＣ（Port Bypass Circuit）７０１を備えている。ファイバチャネルのハードディスクドライブ５１は、このＰＢＣ７０１を介してＦＣ－ＡＬ５０６に接続されている。ＰＢＣ７０１は、チップ化された電子スイッチであり、ディスク制御部５０２やハードディスクドライブ５１などをバイパスし電気的にＦＣ－ＡＬ５０６から除外する機能も有している。具体的には、ＰＢＣ７０１は、障害が発

生したハードディスクドライブ５１をＦＣ－ＡＬ５０６から切り離して、他のハードディスクドライブ５１とディスク制御部５０２との間の通信を可能にする。

【００４５】

また、ＰＢＣ７０１は、ＦＣ－ＡＬ５０６の動作を維持したままでハードディスクドライブ５１の抜き差しを可能にする。例えば、ハードディスクドライブ５１が新たに装着された場合にはそのハードディスクドライブ５１をＦＣ－ＡＬ５０６に取り込み、ディスク制御部５０２との間の通信を可能にする。なお、ＰＢＣ７０１の回路基板は、ディスクアレイ装置１０のラックフレーム１１に設けられているか、もしくは、その一部または全部がコントローラボード５９や電源コントローラボード５６に実装されていることとしてもよい。

【００４６】

図８は、基本筐体２０にシリアルＡＴＡのハードディスクドライブ５１が収容されている場合における、ディスク制御部５０２と各ハードディスクドライブ５１との接続形態を示している。

各ハードディスクドライブ５１は、コンバータ８０１を介してＦＣ－ＡＬ５０６のＰＢＣ７０１に接続されている。コンバータ８０１はファイバチャネルプロトコルとシリアルＡＴＡプロトコルとを変換する回路である。コンバータ８０１は、プロトコル変換機能が組み込まれた１つのチップであり、各ディスクドライブユニット５２内に設けられている。

【００４７】

図９は、基本筐体２０にシリアルＡＴＡのハードディスクドライブ５１が収容されている場合における、もう一つの接続形態を示している。

コンバータ９０１は、図７におけるコンバータ８０１と同様にファイバチャネルプロトコルとシリアルＡＴＡプロトコルとを変換する回路である。コンバータ９０１はＦＣ－ＡＬ５０６のＰＢＣ７０１に接続されており、１つのコンバータ９０１には、複数のハードディスクドライブ５１がスイッチ９０２を介して接続されている。スイッチ９０２は、ハードディスクドライブ５１が複数のコンバータ９０１に接続されている場合において、どのハードディスクドライブ５１と通信を行うかを選択する回路である。スイッチ９０２は、各ディスクドライブユニット５２内に設けられている。コンバータ９０１は、プロトコル変換機能が組み込まれた１つのチップであるか、複数の回路により構成されている。コンバータ９０１は、例えば「米国特許出願公開第２００３／０１３５５７７号明細書」にて開示されているＳＡＴＡマスタデバイスの構成により実現することが可能である。コンバータ９０１は、コントローラボード５９や電源コントローラボード５６等に実装されている。

【００４８】

＝＝信頼性を高めるための制御＝＝

以上に説明したディスクアレイ装置１０において、ハードディスクドライブ５１からの読み出し又はハードディスクドライブ５１への書き込みの信頼性を高める方法について説明する。

【００４９】

＝＝ＲＡＩＤ構成でのパリティチェック＝＝

まず、ＲＡＩＤ構成においてハードディスクドライブ５１に記憶されているデータが不正な状態となっていないか検査する方法について説明する。ここで、不正な状態とは、データがディスク制御部５０２から指定された場所に指定された内容で書き込まれていない状態のことである。

【００５０】

図１０は、ＲＡＩＤ５においてハードディスクドライブ５１にデータが記憶されている様子を表している。ＲＡＩＤ５においては、複数のハードディスクドライブ５１によりＲＡＩＤグループ１００１が形成されている。図１０の例では、ハードディスクドライブ５１には、データＡ～ＤとデータＡ～Ｄに対する誤りを検出するためのパリティデータＰ（

Ａ－Ｄ）が記憶されている。同様に、データＥ～ＨとデータＥ～Ｈに対するパリティデータＰ（Ｅ－Ｈ）が記憶されている。このような、データとパリティデータとの組合せのことを、ストライプグループ１００２と呼ぶこととする。このようなストライプグループ１００２が形成されているＲＡＩＤ構成において、コントローラ５００はストライプグループ１００２の全てのデータ及びパリティデータを読み出すことにより、データが不正な状態となっていないか検査することができる。まず、ＣＰＵ５０３からの指示により、ディスク制御部５０２がデータＡ～Ｄとパリティデータ Ｐ（Ａ－Ｄ）とを読み出す。次に、ＣＰＵ５０３は、データＡ～Ｄとパリティデータ Ｐ（Ａ－Ｄ）とでパリティチェックを行うことにより、データＡ～Ｄのいずれかが不正な状態となっていないか検査することができる。

【００５１】

コントローラ５００は、情報処理装置３００からデータの読み出し要求を受信した際に、当該データを含むストライプグループの全てのデータとパリティデータとを読み出すようにすることもできる。これにより、コントローラ５００がハードディスクドライブ５１から不正なデータを読み出して情報処理装置３００に送信することを防止することが可能となる。なお、不正なデータの検査はデータの読み出し要求を受信した際にかかわらず、任意のタイミングで行うこととしてもよい。これにより、データの読み出し性能に影響を与えずに、不正なデータの検出を行うことが可能となる。

【００５２】

また、図１１に示す更新管理テーブル１１０１を用いて、ハードディスクドライブ５１に書き込まれたデータが不正な状態となっていないか検査することができる。更新管理テーブル１１０１はドライブ番号とセクタ番号とで構成され、メモリ５０４に記憶されている。本実施の形態においてはセクタ番号はＬＢＡ（Logical Block Address）で定義されており、ＬＢＡ＃１－１２８のように１２８ＬＢＡ単位で管理されている。なお、セクタ番号をまとめる単位は１２８に限らず任意の単位でよい。ＣＰＵ５０３は、ディスク制御部５０２を介してデータをハードディスクドライブ５１に書き込むと、更新管理テーブル１１０１の当該ハードディスクドライブ５１の書き込みを行ったセクタの値を「１」に変更する。ＣＰＵ５０３は、更新管理テーブル１１０１において「１」が記憶されているハードディスクドライブ５１の対象セクタのストライプグループの全てのデータとパリティデータとをディスク制御部５０２を介して読み出し、パリティチェックを行う。ＣＰＵ５０３は、読み出したデータが不正でない場合は、更新管理テーブル１１０１において当該セクタの値を「０」に変更する。ＣＰＵ５０３は、チャネル制御部５０１を介して情報処理装置３００からデータの読み出し要求を受信すると、更新管理テーブル１１０１を参照して当該データが記憶されているセクタが検査済みであるかどうかを確認する。当該データが記憶されているセクタが検査済みでない場合は、ＣＰＵ５０３は前述の手順に従い当該データの属するストライプグループのデータを検査する。このように、ハードディスクドライブ５１に書き込まれたデータについて、当該データに対する読み出し要求を受信する前に当該データの検査を実施しておくことにより、データの読み出し性能の低下を抑えることが可能となる。また、検査の未済を更新管理テーブル１１０１に記憶し、検査が行われていないデータを読み出す際にはパリティチェックを行うため、不正データの読み出しを防止することが可能となる。

【００５３】

＝＝ＷＲＩＴＥデータに対する検査＝＝

次に、ハードディスクドライブ５１にデータを書き込んだ際に、当該データが正しく書き込まれているか検査する方法について説明する。

【００５４】

図１２は、コントローラ５００がハードディスクドライブ５１にデータを書き込む際のＣＰＵ５０３での制御を表すフローチャートである。ＣＰＵ５０３は、チャネル制御部５０１を介して情報処理装置３００からデータの書き込み要求を受信すると、当該データのハードディスクドライブ５１への書き込み指示をディスク制御部５０２に送信する（Ｓ１

２０１）。そして、ＣＰＵ５０３は当該データが書き込まれた磁気ディスクのヘッドの位置を移動させるシーク処理の実行指示をディスク制御部に送信する（Ｓ１２０２）。次に、ＣＰＵ５０３は、キャッシュメモリ６２から当該データを読み出し（Ｓ１２０３）、磁気ディスクから当該データを読み出す（Ｓ１２０４）。ＣＰＵ５０３は、キャッシュメモリ６２のデータと磁気ディスクのデータとが一致しているか比較する（Ｓ１２０５）。２つのデータが一致していない場合、ＣＰＵ５０３は、書き込みが正常に行われていないことを情報処理装置３００に通知する（Ｓ１２０６）。

このように、磁気ディスクに記憶されているデータとキャッシュメモリ６２に記憶されているデータとを比較することにより、磁気ディスクに正しくデータが書き込まれているか確認することが可能である。また、書き込まれているデータが不正な状態となっている場合においても、データがキャッシュメモリ６２に残っているため、データを失うことがない。なお、比較するデータを磁気ディスクとキャッシュメモリ６２とから読み出す前に、シーク処理等により当該ハードディスクドライブが備えるヘッドを移動することにより、書き込み時にヘッドの位置が不正であった場合に、同じ位置から再度読み出すことを防止することが可能となる。

【００５５】

図１２の処理においては、書き込まれたデータの全部をキャッシュメモリ６２と磁気ディスクとから読み出し比較することによりデータの検査を実施したが、データの全部ではなく、先頭と末尾の１セグメント等、そのデータの一部を読み出して比較することとしてもよい。例えば、シリアルＡＴＡのハードディスクドライブは、データのバックアップ等の用途に用いられるため、サイズの大きいデータ（シーケンシャルデータ）が書き込まれることが多い。このような場合、書き込みを行ったデータの全部について、磁気ディスクに記憶されているデータとキャッシュメモリ６２に記憶されているデータとを比較することは、書き込み処理の性能を著しく低下させる要因となる。また、シーケンシャルデータの書き込み時に書き込み位置の誤り等が発生した場合は、そのデータの全部が不正となっている可能性が高い。そのため、データの一部を検査することでデータが不正となっているか判断可能である場合が多い。つまり、書き込みを行ったデータの一部、例えば先頭と末尾の１セグメント等について比較することにより、書き込み処理の性能低下を抑えたうえで、不正データのチェックを行うことが可能である。

【００５６】

また、ハードディスクドライブ５１に書き込まれたデータのサイズに応じて、データの検査方法を変更することとしてもよい。図１３は、書き込まれたデータがシーケンシャルデータであるかどうかに応じて、検査方法を変更する処理を示すフローチャートである。ＣＰＵ５０３は、ハードディスクドライブ５１にデータを書き込む指示をディスク制御部５０２に送信する（Ｓ１３０１）。そして、ＣＰＵ５０３は当該データが書き込まれた磁気ディスクのヘッドの位置を移動させるシーク処理の実行指示をディスク制御部に送信する（Ｓ１３０２）。ＣＰＵ５０３は、当該データがシーケンシャルデータであるかどうか判断する（Ｓ１３０３）。なお、シーケンシャルデータであるかどうかの判断は、書き込まれたデータのサイズが既定のサイズ以上であるかどうかにより行う。

ＣＰＵ５０３は、当該データがシーケンシャルデータである場合は、キャッシュメモリ６２と磁気ディスクとから当該データの先頭と末尾の１セグメントを読み出す。また、ＣＰＵ５０３は、当該データがシーケンシャルデータでない場合は、キャッシュメモリ６２と磁気ディスクとから当該データの全部を読み出す（Ｓ１３０６，Ｓ１３０７）。その後、ＣＰＵ５０３は読み出した２つのデータが一致しているか比較し（Ｓ１３０８）、一致していない場合は書き込みが正常に行われていないことを情報処理装置３００に通知する（Ｓ１３０９）。

このように、書き込んだデータがシーケンシャルデータである場合は、当該データの一部について、磁気ディスクとキャッシュメモリ６２とに記憶されているデータを比較することにより、書き込み処理の性能低下を抑えた上で、データの不正を検出することが可能である。また、書き込んだデータがシーケンシャルデータでない場合は、書き込んだデー

タの全部について、磁気ディスクとキャッシュメモリ６２とに記憶されているデータを比較することにより、シーケンシャルデータの場合ほど書き込み処理の性能を著しく低下させることなく、データの不正を完全に検出することが可能である。

【００５７】

ハードディスクドライブ５１は、データの書き込み性能を向上させるため、コントローラ５００からデータの書き込み要求を受信すると、当該データをディスクキャッシュのみに書き込み、コントローラ５００に書き込み完了を通知する機能を備えている場合がある。この場合、図１２および図１３に説明した方法では、書き込まれたデータの検査を行うことができない。図１４は、ハードディスクドライブ５１がこのような機能を備えている場合において、書き込まれたデータの検査を行う処理のフローチャートを示す図である。ＣＰＵ５０３は、ハードディスクドライブ５１への書き込み回数が所定の回数を超過していないか監視している（Ｓ１４０１）。所定の回数を超過すると、ＣＰＵ５０３はディスクキャッシュに記憶されているデータを磁気ディスクに書き込むよう、ディスク制御部４５０２を介してハードディスクドライブに通知する（Ｓ１４０２）。そして、ＣＰＵ５０３は、当該データをキャッシュメモリ６２と磁気ディスクとから読み出す（Ｓ１４０３，Ｓ１４０４）。ＣＰＵ５０３は、キャッシュメモリ６２のデータと磁気ディスクのデータとが一致しているか確認し（Ｓ１４０５）、一致していない場合は書き込みが正常に行われていないことを情報処理措置３００に通知する（Ｓ１４０６）。これにより、前述した書き込み処理の性能を高める機能を用いた上で、データの不正を検出することが可能となる。なお、図１４の処理では、書き込み回数が所定の回数を超過した場合に磁気ディスクへのデータの書き込みと書き込まれたデータの検査を行うこととしたが、所定の時間が経過した場合や、ディスクキャッシュの空領域が無くなった場合などを契機としてもよい。

【００５８】

また、シリアルＡＴＡのハードディスクドライブ５１においては、ヘッドの障害により、データの書き込みが正しく行われていないことが多い。そこで、ハードディスクドライブ５１からデータを読み出す際に、ヘッドの障害を検出する方法を説明する。

【００５９】

図１５はヘッドチェック管理テーブル１５０１を示す図である。ヘッドチェック管理テーブル１５０１はドライブ番号とヘッド番号とセクタ番号とで構成され、メモリ５０４に記憶されている。セクタ番号は更新管理テーブル１１０１と同様にＬＢＡにより定義されている。ＣＰＵ５０３は、ディスク制御部５０２を介してデータをハードディスクドライブ５１に書き込むと、ヘッドチェック管理テーブル１５０１の当該データの書き込みを行ったヘッドの当該セクタの「更新有無」の値を「１」に変更する。

【００６０】

図１６は、ＣＰＵ５０３が実行するヘッドチェック処理のフローチャートを示す図である。ＣＰＵ５０３は、検査ヘッド番号に初期値として１を設定する（Ｓ１６０１）。ＣＰＵ５０３は、一定時間経過するのを待ち（Ｓ１６０２）、検査ヘッド番号で指定されるヘッドを用いて磁気ディスクの管理領域に検査用のデータを書き込む（Ｓ１６０３）。なお管理領域は磁気ディスク上の予め定められている記憶領域である。次にＣＰＵ５０３は、管理領域に書き込まれているデータを読み出し（Ｓ１６０４）、読み出したデータと検査用のデータとが一致しているか確認する（Ｓ１６０５）。

【００６１】

データが一致している場合、ＣＰＵ５０３は当該ヘッドに異常が無いと判断し、ヘッドチェック管理テーブル１５０１の当該ヘッドの「更新有無」を「０」に変更する（Ｓ１６０６）。ＣＰＵ５０３は、検査ヘッド番号に１を加算する（Ｓ１６０７）。検査ヘッド番号がヘッド番号の最大値より大きいか確認し（Ｓ１６０８）、大きい場合は検査ヘッド番号を１に設定する。ＣＰＵ５０３は、設定されたヘッド番号について、ヘッドチェック処理を繰り返し実行する。

【００６２】

管理領域から読み出したデータと検査用のデータとが一致していない場合、ＣＰＵ５０

３は当該ハードディスクドライブ５１に異常が発生していることを情報処理装置３００に通知し、処理を終了する。

【００６３】

図１７は、ＣＰＵ５０３が情報処理装置３００からデータの読み出し要求を受信した際の処理のフローチャートを示す図である。ＣＰＵ５０３は、チャネル制御部５０１を介して情報処理装置３００からデータの読み出し要求を受信する（Ｓ１７０１）。ＣＰＵ５０３は、ヘッドチェック管理テーブル１５０１から、当該データが記憶されているハードディスクドライブ５１の対象セクタの「更新有無」を確認する（Ｓ１７０２，Ｓ１７０３）。「更新有無」が「１」である場合は、当該ハードディスクドライブ５１の当該ＬＢＡについてデータの書き込みが行われているが、前述したヘッドチェック処理が行われていない状態を示している。「更新有無」が「０」である場合は、ＣＰＵ５０３は当該データをハードディスクドライブ５１から読み出す（Ｓ１７０８）。

【００６４】

「更新有無」が「１」である場合、ＣＰＵ５０３は前述したヘッドチェック処理と同様に、当該ヘッドを用いて磁気ディスクの管理領域に検査用のデータを書き込む（Ｓ１７０４）。なお管理領域は磁気ディスク上の予め定められている記憶領域である。次にＣＰＵ５０３は、管理領域に書き込まれているデータを読み出し（Ｓ１７０５）、読み出したデータと検査用のデータとが一致しているか確認する（Ｓ１７０６）。

データが一致している場合、ＣＰＵ５０３は当該ヘッドに異常が無いと判断し、ヘッドチェック管理テーブル１５０１の当該ヘッドの「更新有無」を「０」に変更する（Ｓ１７０７）。そして、ＣＰＵ５０３は当該読み出し要求に従いハードディスクドライブ５１からデータを読み出す（Ｓ１７０８）。

管理領域から読み出したデータと検査用のデータとが一致していない場合、ＣＰＵ５０３は当該ハードディスクドライブ５１に異常が発生していることを情報処理装置３００に通知し（Ｓ１７０９）、当該ハードディスクドライブ５１からデータを読み出さずに処理を終了する。

【００６５】

これにより、ハードディスクドライブ５１に書き込まれているデータを読み出す際に、当該データの書き込みを行ったヘッドが正常であるかどうかを確認することができる。ヘッドが異常である場合、データが正しく書き込まれていない可能性や、データの読み出しを正しく行うことができない可能性がある。データの読み出し時にヘッドの異常を検知することにより、不正なデータを読み出すことを防止することが可能となる。

【００６６】

＝＝パリティ付与による検査＝＝

前述したＲＡＩＤ構成でのストライプグループの全てのデータを読み出してパリティチェックする方法では、ストライプグループの中のどのデータが不正な状態となっているか判断することができなかった。そのため、不正なデータの読み出しを防止することは可能であるが、不正なデータを復旧させることができず、データを損失してしまう可能性がある。そこで、ストライプグループにおけるパリティデータとは別に、各データにパリティデータを付与する方法について説明する。

【００６７】

ＣＰＵ５０３は、データをハードディスクドライブ５１に書き込む際の最小単位を必ず複数セクタとし、これら複数セクタに対する誤りを検出するためのパリティデータを生成する。本実施の形態において、この複数セクタのデータとパリティデータとの組合せをデータユニットと称することとする。ＣＰＵ５０３は、チャネル制御部５０１を介して情報処理装置３００からデータの書き込み要求を受信すると、当該データからデータユニットを形成する。ＣＰＵ５０３は当該データユニットをディスク制御部５０２を介してハードディスクドライブ５１に書き込む。

図１８は、ハードディスクドライブに１つのデータ１８０１が書き込まれている様子を示す図である。データ１８０１は複数のセクタＳ＃１～Ｓ＃４で構成され、これら複数セ

クタのデータ１８０１に対するパリティデータ１８０２とでデータユニット１８０３が形成されている。ＣＰＵ５０３はチャネル制御部５０１を介して情報処理装置３００からデータの読み出し要求を受信すると、ディスク制御部５０２を介して当該データのデータユニット１８０３を読み出し、当該データのパリティチェックを行うことで当該データが不正な状態となっていないか検査する。このように、読み出し要求の対象となっているデータのみを読み出して、当該データが不正な状態となっていないか判断することが可能となる。また、ハードディスクドライブ５１がＲＡＩＤ５のように冗長性のあるＲＡＩＤ構成である場合には、ストライプグループにおける他のデータ及びパリティデータとを用いて、当該データを復元することが可能であるため、データを損失することがない。

【００６８】

１つのハードディスクドライブ５１においてヘッドの障害等が発生している場合は、不正な状態となっているセクタが複数発生する可能性が高い。データユニット１８０３が１つのハードディスクドライブ５１に書き込まれている場合にデータユニット１８０３のうちの複数のセクタが不正な状態となると、パリティチェックにより不正を検出できない場合がある。

【００６９】

そこで、図１９に示すように、ＣＰＵ５０３は、前述したデータユニット１８０３をディスク制御部５０２を介してＲＡＩＤグループ内の複数のハードディスクドライブ５１に分散して書き込むこととしてもよい。図２０は、データユニット管理テーブル２００１を示している図である。データユニット管理テーブル２００１は、複数セクタで構成されるデータユニット１８０３がどのハードディスクドライブ５１のどのＬＢＡに対応しているかを示している。図２０の例では、０００～１２９までの１３０セクタで１つのデータユニット１８０３が形成され、このデータユニットはドライブ番号＃０とドライブ番号＃１のハードディスクドライブ５１の０００～０６４までのＬＢＡにより構成されていることが示されている。ＣＰＵ５０３は、情報処理装置３００からデータの書き込み要求を受信すると、データユニット管理テーブル２００１を参照し、データユニット１８０３ごとに複数のハードディスクドライブ５１に分散してデータを書き込む。

【００７０】

これにより、１つのハードディスクドライブに障害が発生している場合においても、データの不正を検出することができる可能性が高くなる。

【００７１】

＝＝ファイバチャネルとシリアルＡＴＡとが混在する環境＝＝

次に、ファイバチャネルのハードディスクドライブ５１とシリアルＡＴＡのハードディスクドライブ５１とが混在しているディスクアレイ装置１０における説明を行う。

【００７２】

図２１は、第一の筐体２１０１にファイバチャネルのハードディスクドライブ５１、第二の筐体２１０２にシリアルＡＴＡのハードディスクドライブ５１が収容されているディスクアレイ装置を示すブロック図である。なお、第一の筐体２１０１及び第二の筐体２１０２とは、基本筐体２０または増設筐体３０のことである。各ハードディスクドライブ５１のディスク制御部５０２との接続形態については、前述した通りである。また、図１９においては、１つのコンバータ９０１に複数のシリアルＡＴＡのハードディスクドライブが接続される形態を示しているが、前述したように、コンバータ８０１が各ディスクドライブユニットに設けられて接続されているものとしてもよい。

【００７３】

このようなディスクアレイ装置１０においては、ファイバチャネルのハードディスクドライブ５１と比較して信頼性の低いシリアルＡＴＡのハードディスクドライブ５１の信頼性を高めることが求められている。そこで、コントローラ５００は、シリアルＡＴＡのハードディスクドライブ５１のみに対して、前述した信頼性を高める方法を適用する。これにより、基幹業務等の高いアクセス性能が要求される処理に用いられるファイバチャネルのハードディスクドライブ５１に対するデータの読み書き性能を落とさずに、シリアルＡ

ＴＡのハードディスクドライブ５１に対するデータの読み書きの信頼性を高めることができる。また、シリアルＡＴＡのハードディスクドライブ５１の各磁気ディスクにヘッドを２つずつ設ける等の物理的な構造の変更が必要でないため、シリアルＡＴＡのハードディスクドライブ５１の製造コストを抑えることが可能である。

【００７４】

なお、本実施の形態においては、ファイバチャネルのハードディスクドライブ５１とシリアルＡＴＡのハードディスクドライブ５１とが混在しているとしたが、信頼性の異なるインタフェース規格のハードディスクドライブ５１であれば他のものでもよい。例えば、シリアルＡＴＡのハードディスクドライブ５１の代わりにパラレルＡＴＡのハードディスクドライブ５１であるとしてもよい。

【００７５】

以上、本実施の形態について説明したが、上記実施例は本発明の理解を容易にするためのものであり、本発明を限定して解釈するためのものではない。本発明は、その趣旨を逸脱することなく、変更、改良され得ると共に、本発明にはその等価物も含まれる。

【図面の簡単な説明】

【００７６】

【図１】本実施の形態における、ディスクアレイ装置の外観を示す図である。

【図２】本実施の形態における、ディスクアレイ装置の基本筐体の構成を示す図である。

【図３】本実施の形態における、ディスクアレイ装置の増設筐体の構成を示す図である。

【図４】本実施の形態における、ハードディスクドライブの構成を示す図である。

【図５】本実施の形態における、ディスクアレイ装置の構成を示す図である。

【図６】本実施の形態における、コントローラのＣＰＵが実行するマイクロプログラムがメモリに記憶されている状態を示す図である。

【図７】本実施の形態における、ファイバチャネルのハードディスクドライブをコントローラのディスク制御部と接続する形態を示す図である。

【図８】本実施の形態における、シリアルＡＴＡのハードディスクドライブをコントローラのディスク制御部と接続する第一の形態を示す図である。

【図９】本実施の形態における、シリアルＡＴＡのハードディスクドライブをコントローラのディスク制御部と接続する第二の形態を示す図である。

【図１０】本実施の形態における、ＲＡＩＤグループを構成するハードディスクドライブにデータが書き込まれている例を示す図である。

【図１１】本実施の形態における、更新管理テーブルを示す図である。

【図１２】本実施の形態における、データ書き込み時にキャッシュメモリと磁気ディスクとに記憶されているデータを比較するフローチャートを示す図である。

【図１３】本実施の形態における、データ書き込み時にデータサイズを考慮してキャッシュメモリと磁気ディスクとに記憶されているデータを比較するフローチャートを示す図である。

【図１４】本実施の形態における、ディスクキャッシュに記憶されているデータを磁気ディスクに書き込む際にキャッシュメモリと磁気ディスクとに記憶されているデータを比較するフローチャートを示す図である。

【図１５】本実施の形態における、ヘッドチェック管理テーブルを示す図である。

【図１６】本実施の形態における、定期的に実施するヘッドチェックのフローチャートを示す図である。

【図１７】本実施の形態における、データの読み出し時にヘッドチェックを実施するフローチャートを示す図である。

【図１８】本実施の形態における、データユニットが１つのハードディスクドライブに書き込まれている例を示す図である。

【図１９】本実施の形態における、データユニットが複数のハードディスクドライブに分散して書き込まれている例を示す図である。

【図２０】本実施の形態における、データユニット管理テーブルを示す図である。

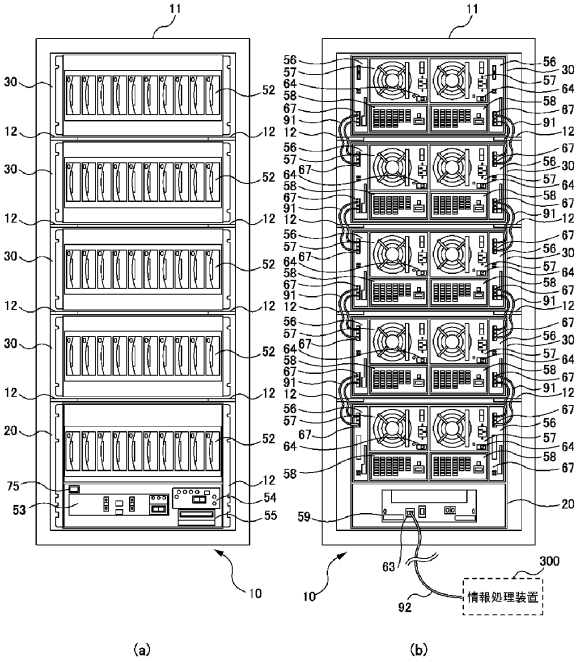
【図２１】本実施の形態における、第一の筐体にファイバチャネルのハードディスクドライブが収容され、第二の筐体にシリアルＡＴＡのハードディスクドライブが収容されているディスクアレイ装置の構成を示す図である。

【符号の説明】

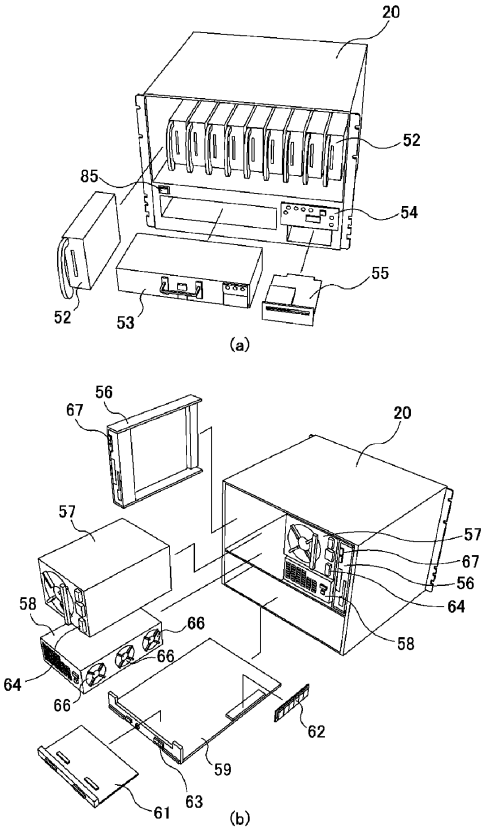
【００７７】

| | |
|---|---|
| １０　ディスクアレイ装置 | １１　ラックフレーム |
| １２　マウントフレーム | ２０　基本筐体 |
| ３０　増設筐体 | ４８　制御ライン |
| ４９　電源供給ライン | ５１　ハードディスクドライブ |
| ５２　ディスクドライブユニット | ５６　電源コントローラボード |
| ５７　ＡＣ／ＤＣ電源 | ５８　冷却ファンユニット |
| ５９　コントローラボード | ６１　通信インターフェースボード |
| ６２　キャッシュメモリ | ６４　ブレーカスイッチ |
| ６６　冷却ファン | ６７　コネクタ |
| ７０　ディスクドライブの筐体 | ７３　磁気ディスク |
| ８５　メインスイッチ | ８１　電源コントローラ |
| ９１　ファイバチャネルケーブル | ３００　情報処理装置 |
| ５００　コントローラ | ５０１　チャネル制御部 |
| ５０２　ディスク制御部 | ５０３　ＣＰＵ |
| ５０４　メモリ | ５０５　データコントローラ |
| ５０６　ＦＣ－ＡＬ | ７０１　ＰＢＣ |
| ８０１　コンバータ | ９０１　コンバータ |
| ９０２　スイッチ | １００１　ＲＡＩＤグループ |
| １００２　ストライプグループ | １１０１　更新管理テーブル |
| １５０１　ヘッドチェック管理テーブル | |
| １８０１　複数セクタのデータ | １８０２　パリティデータ |
| １８０３　データユニット | ２００１　データユニット管理テーブル |

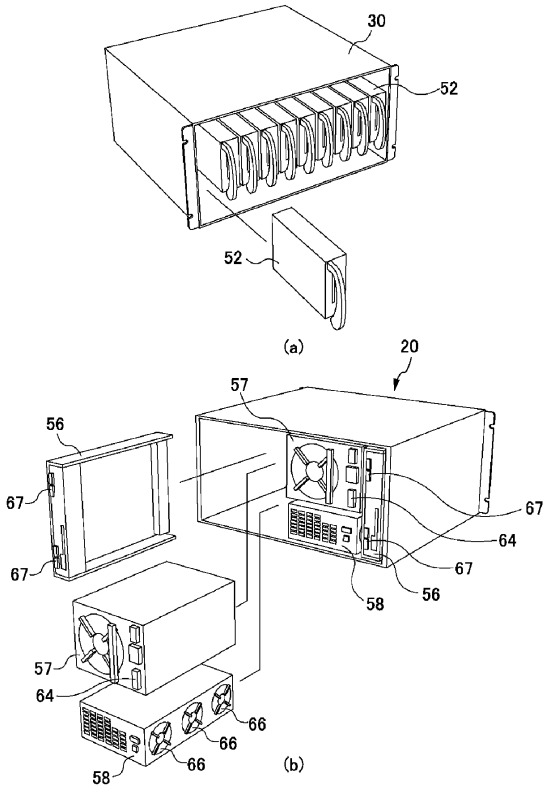
【図１】


11
30
52
12
20
75
53
54
55
10
(a)
(b)
56
57
64
58
67
91
59
63
92
300
情報処理装置

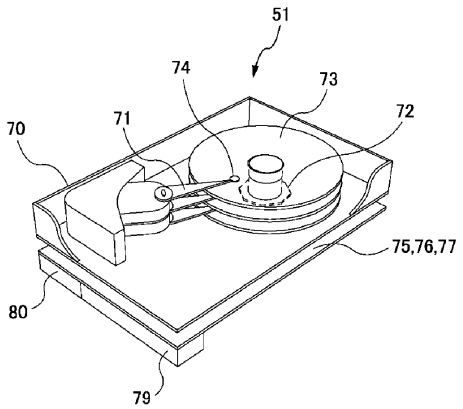
【図２】


20
52
85
54
55
53
(a)
56
67
57
58
66
64
62
59
61
63
(b)

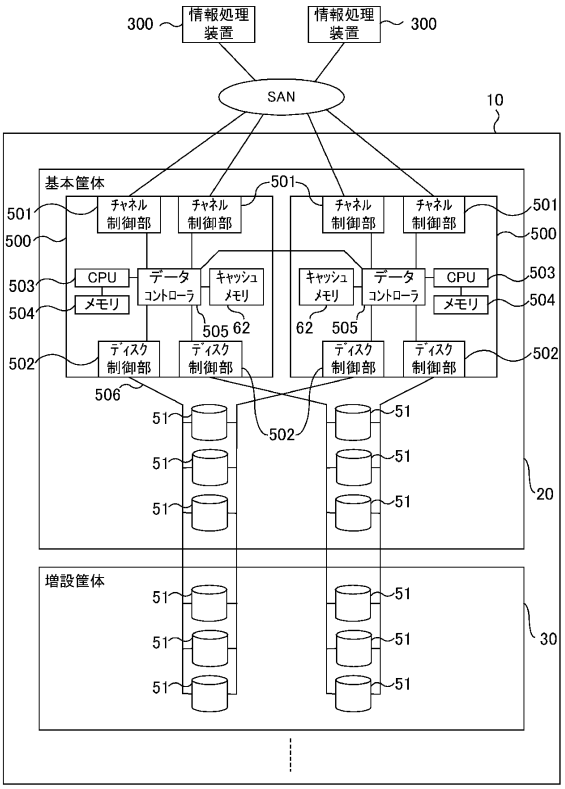
【図３】


30
52
(a)
20
57
56
67
64
58
66
(b)

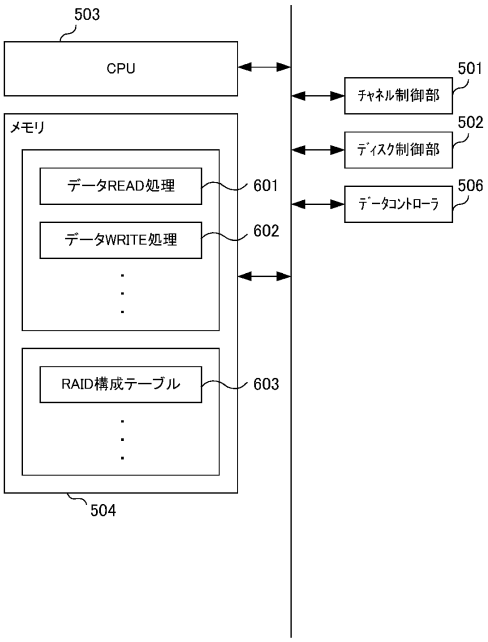
【図４】


51
74
73
71
72
70
75,76,77
80
79

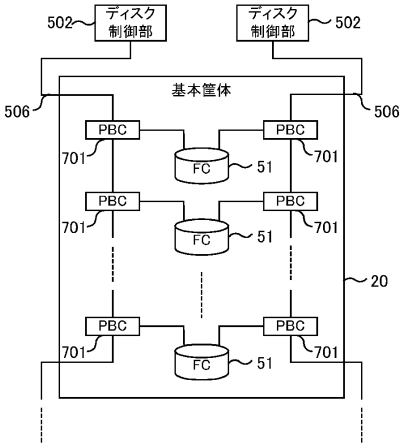
【図５】


300
情報処理装置
情報処理装置
300
SAN
10
基本筐体
501
チャネル制御部
チャネル制御部
チャネル制御部
チャネル制御部
501
500
CPU
データコントローラ
キャッシュメモリ
キャッシュメモリ
データコントローラ
CPU
503
メモリ
メモリ
504
505
62
62
505
ディスク制御部
ディスク制御部
ディスク制御部
ディスク制御部
502
506
51
20
増設筐体
30

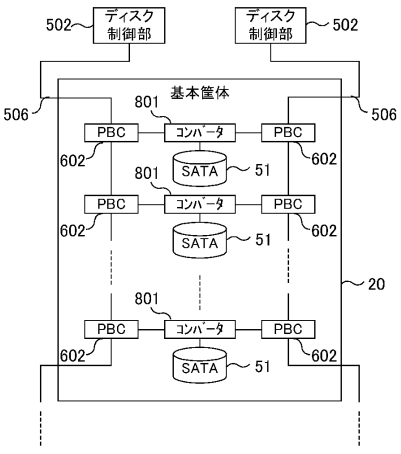
【図６】


503
CPU
501
チャネル制御部
502
ディスク制御部
506
データコントローラ
メモリ
データREAD処理
601
データWRITE処理
602
RAID構成テーブル
603
504

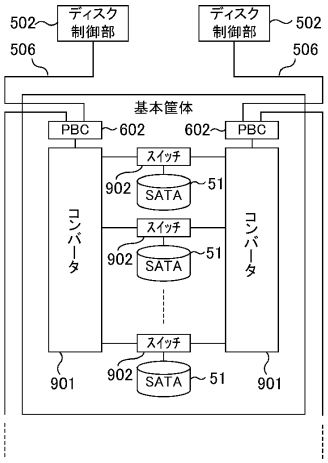
【図７】


502
ディスク制御部
ディスク制御部
502
基本筐体
506
506
PBC
PBC
701
701
FC
51
PBC
PBC
701
701
FC
51
20
PBC
PBC
701
701
FC
51

【図８】


502
ディスク制御部
ディスク制御部
502
基本筐体
801
506
506
PBC
コンバータ
PBC
602
602
SATA
51
801
PBC
コンバータ
PBC
602
602
SATA
51
20
801
PBC
コンバータ
PBC
602
602
SATA
51

【図９】

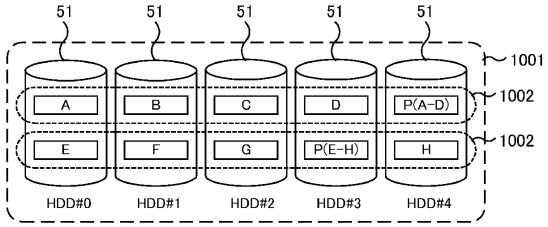



【図１１】

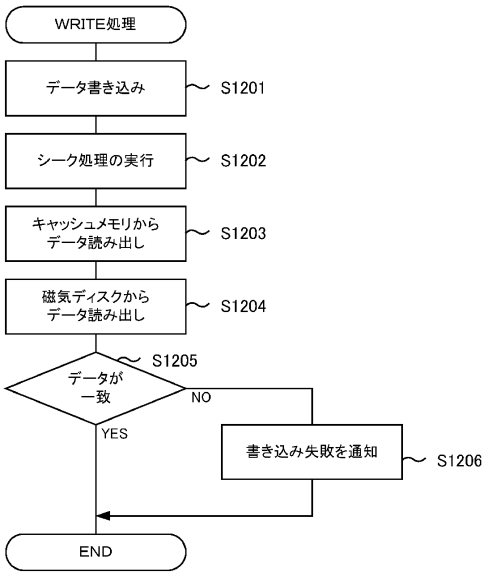
| セクタ番号 / ドライブ番号 | LBA #1-128 | LBA #129-256 | LBA #257-384 | … |
|---|---|---|---|---|
| HDD#0 | 0 | 0 | 1 | … |
| HDD#1 | 1 | 0 | 0 | … |
| HDD#2 | 0 | 1 | 0 | … |
| … | … | … | … | … |

1101

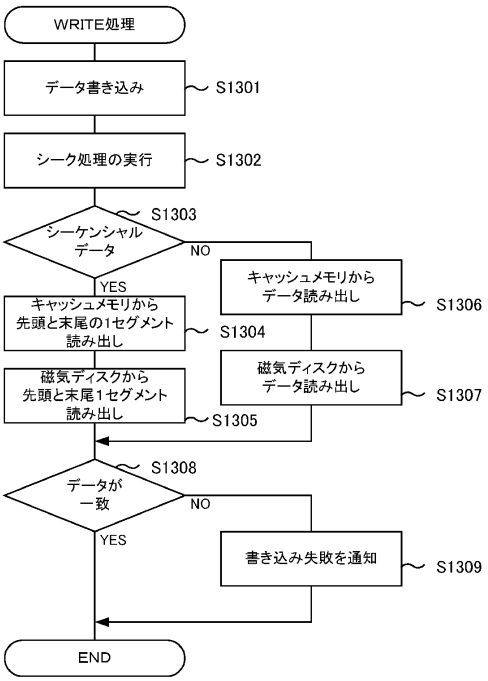
【図１０】



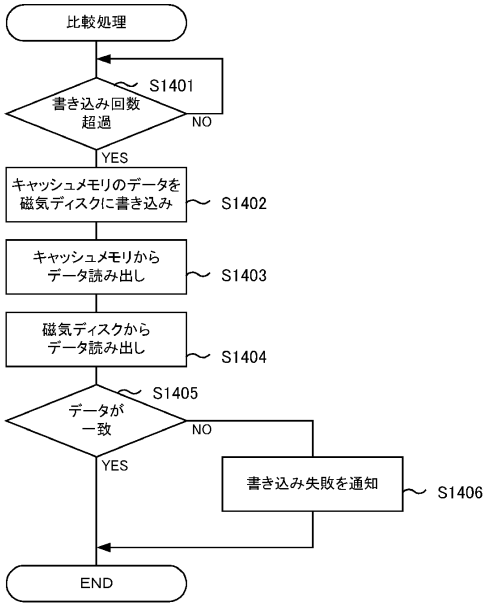

【図１２】




【図１３】



【図１４】

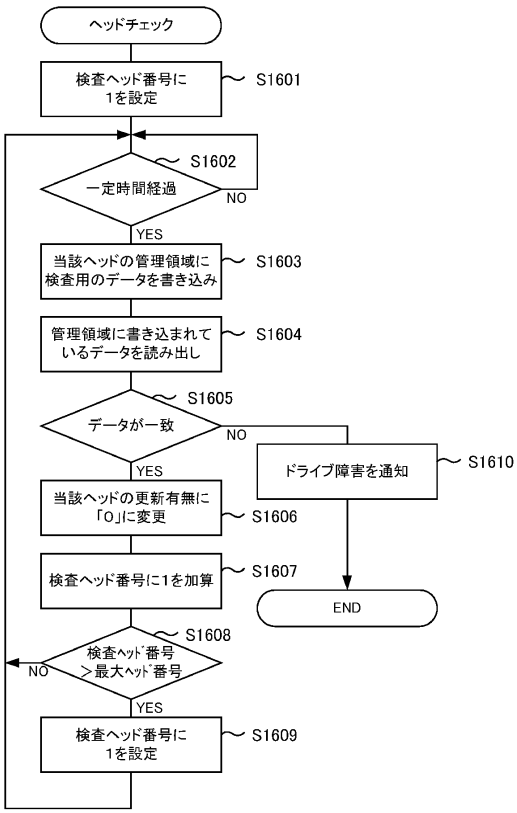



【図１５】

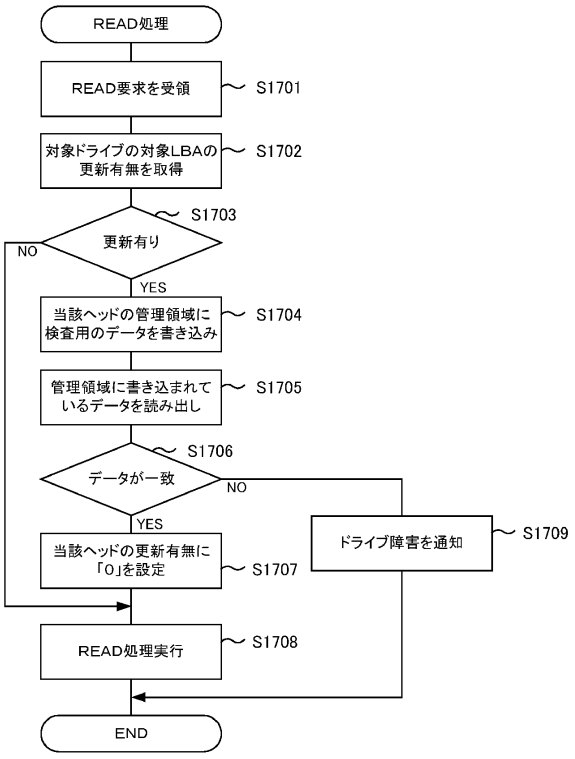
| | | LBA #1-128 | LBA #129-256 | LBA #257-384 | … |
|---|---|---|---|---|---|
| HDD#0 | ﾍｯﾄﾞ番号 | #0 | #0 | #1 | … |
| | 更新有無 | 1 | 0 | 0 | … |
| HDD#1 | ﾍｯﾄﾞ番号 | #0 | #0 | #1 | … |
| | 更新有無 | 0 | 1 | 0 | … |
| … | … | … | … | … | … |

1501

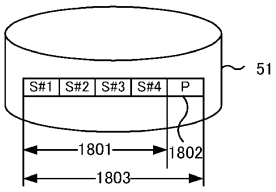
【図１６】



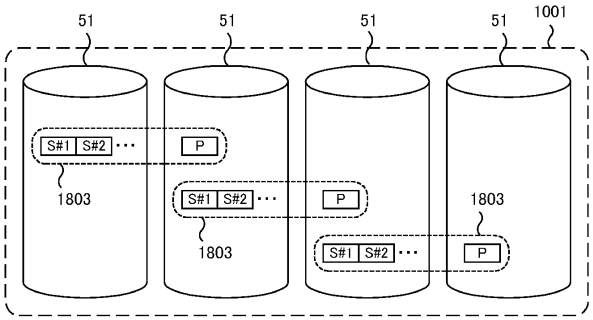

【図１７】



【図１８】




【図１９】

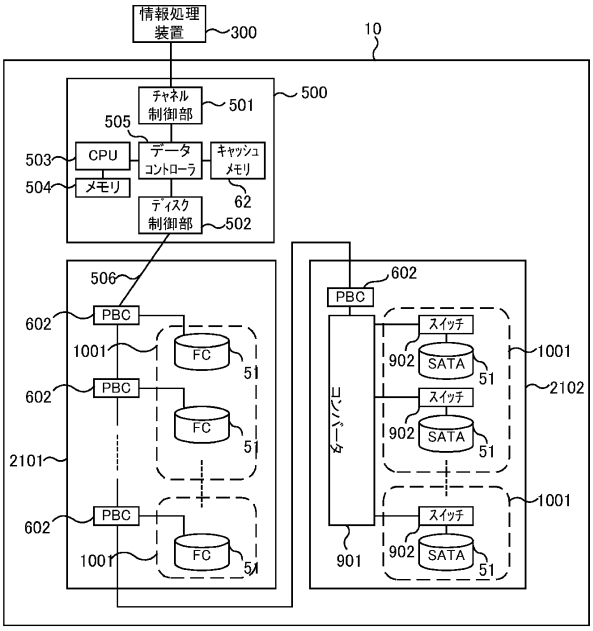



【図２０】

| データユニット | ドライブ | ドライブLBA |
|---|---|---|
| 000-129 | #0 | 000-064 |
| | #1 | 000-064 |
| 130-259 | #2 | 000-064 |
| | #0 | 065-129 |
| ⋮ | ⋮ | ⋮ |

2001

【図２１】



フロントページの続き

(72)発明者　八木沢　育哉
神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所　システム開発研究所内

審査官　藤井　浩

(56)参考文献　特開２００２－００７０７７（ＪＰ，Ａ）
国際公開第０３／０８３６３６（WO，Ａ１）
特表２００５－５２０２５８（ＪＰ，Ａ）
特開２００３－１０８５０８（ＪＰ，Ａ）
特開２００３－１０８５１０（ＪＰ，Ａ）
特開２００１－２２２３８５（ＪＰ，Ａ）
特開２００３－３０３０５５（ＪＰ，Ａ）
米国特許出願公開第２００３／０１３５５７７（ＵＳ，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　　３／０６　－　３／０８