

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4651913号
(P4651913)

(45) 発行日 平成23年3月16日(2011.3.16)

(24) 登録日 平成22年12月24日(2010.12.24)

(51) Int.Cl. F I
G 0 6 F 3/06 (2006.01) G O 6 F 3/06 3 O 4 P
G 1 1 B 27/00 (2006.01) G O 6 F 3/06 5 4 O
 G 1 1 B 27/00 D

請求項の数 13 (全 28 頁)

<p>(21) 出願番号 特願2003-37602(P2003-37602) (22) 出願日 平成15年2月17日(2003.2.17) (65) 公開番号 特開2004-246749(P2004-246749A) (43) 公開日 平成16年9月2日(2004.9.2) 審査請求日 平成18年2月16日(2006.2.16) 審判番号 不服2008-18202(P2008-18202/J1) 審判請求日 平成20年7月17日(2008.7.17)</p> <p>早期審査対象出願</p>	<p>(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 110000062 特許業務法人第一国際特許事務所 (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 (72) 発明者 松並 直人 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</p>
---	--

最終頁に続く

(54) 【発明の名称】 記憶装置システム

(57) 【特許請求の範囲】

【請求項1】

複数のディスク装置と、
 前記複数のディスク装置に接続される制御部とを有し、
 前記複数のディスク装置には、各々が第一のタイプのインタフェースを有する複数の第一のインタフェースタイプのディスク装置と、各々が第二のタイプのインタフェースを有する複数の第二のインタフェースタイプのディスク装置とが含まれる記憶装置システムであって、
 前記制御部は、前記複数のディスク装置のインタフェースタイプに関する情報を取得し、
 前記制御部は、取得した前記複数のディスク装置のインタフェースタイプに関する情報を出力し、
 前記制御部は、前記複数のディスク装置の中から、前記出力された前記複数のディスク装置のインタフェースタイプに関する情報に基づいて選択されたディスク装置でRAIDグループを作成するよう指示を受け、
 前記制御部は、前記複数の第一のインタフェースタイプのディスク装置を用いて少なくとも一の第一のインタフェースタイプのRAIDグループを、前記複数の第二のインタフェースタイプのディスク装置を用いて前記第一のインタフェースタイプのRAIDグループとは異なる少なくとも一の第二のインタフェースタイプのRAIDグループを作成する前記指示に基づいて、前記選択されたディスク装置を用いて前記複数の第一のインタフェ

ースタイプのディスク装置を用いて少なくとも一の第一のインタフェースタイプのRAIDグループを、前記複数の第二のインタフェースタイプのディスク装置を用いて前記第一のインタフェースタイプのRAIDグループとは異なる少なくとも一の第二のインタフェースタイプのRAIDグループを作成することを特徴とする記憶装置システム。

【請求項2】

請求項1記載の記憶装置システムであって、

前記第一のタイプのインタフェースはFCインタフェース、前記第二のタイプのインタフェースはATAインタフェースであることを特徴とする記憶装置システム。

【請求項3】

請求項1又は2記載の記憶装置システムであって、

前記第一のインタフェースタイプのRAIDグループには第一のインタフェースタイプの論理記憶領域が、前記第二のインタフェースタイプのRAIDグループには第二のインタフェースタイプの論理記憶領域が構成され、

前記制御部は各論理記憶領域が第一のインタフェースタイプの論理記憶領域若しくは第二のインタフェースタイプの論理記憶領域のいずれであるかを管理しており、

前記制御部が管理する論理記憶領域のインタフェースタイプに基づいて、前記第一のインタフェースタイプの論理記憶領域に格納されるデータのコピーを格納する論理記憶領域として、前記第二のインタフェースタイプの論理記憶領域が選択されることを特徴とする記憶装置システム。

【請求項4】

請求項3記載の記憶装置システムであって、

前記制御部は、前記第一のインタフェースタイプの論理記憶領域から前記第二のインタフェースタイプの論理記憶領域へのデータのコピー処理を実行することを特徴とする記憶装置システム。

【請求項5】

請求項1又は2記載の記憶装置システムであって、

前記制御部は、前記複数のディスク装置に対してコマンドを発行し、当該コマンドの応答として、前記複数のディスク装置から当該複数のディスク装置のインタフェースタイプに関する情報を取得することを特徴とする記憶装置システム。

【請求項6】

複数のディスク装置と、

前記複数のディスク装置に接続される制御部とを有し、

前記複数のディスク装置には、各々が第一のタイプのインタフェースを有する複数の第一のインタフェースタイプのディスク装置と、各々が第二のタイプのインタフェースを有する複数の第二のインタフェースタイプのディスク装置とが含まれる記憶装置システムにおいて、当該制御部によって実行される方法であって、

前記制御部は、前記複数のディスク装置のインタフェースタイプの情報を取得し、

前記制御部は、取得した前記複数のディスク装置のインタフェースタイプの情報を出力し、

前記制御部は、前記複数のディスク装置の中から出力された前記複数のディスク装置のインタフェースタイプの情報を用いて選択されたディスク装置でRAIDグループを作成するよう指示を受け、

前記制御部は、前記複数の第一のインタフェースタイプのディスク装置を用いて少なくとも一の第一のインタフェースタイプのRAIDグループを、前記複数の第二のインタフェースタイプのディスク装置を用いて少なくとも一の第二のインタフェースタイプのRAIDグループを作成する前記指示に基づいて、前記選択されたディスク装置を用いて前記複数の第一のインタフェースタイプのディスク装置を用いて少なくとも一の第一のインタフェースタイプのRAIDグループを、前記複数の第二のインタフェースタイプのディスク装置を用いて少なくとも一の第二のインタフェースタイプのRAIDグループを作成することを特徴とする方法。

10

20

30

40

50

【請求項 7】

請求項 6 記載の方法であって、

前記第一のタイプのインタフェースは F C インタフェース、前記第二のタイプのインタフェースは A T A インタフェースであることを特徴とする方法。

【請求項 8】

請求項 6 又は 7 記載の方法であって、

前記第一のインタフェースタイプの R A I D グループには第一のインタフェースタイプの論理記憶領域が、前記第二のインタフェースタイプの R A I D グループには第二のインタフェースタイプの論理記憶領域が構成されており、

前記制御部は、各論理記憶領域のインタフェースタイプを管理し、

前記制御部は、前記第一のインタフェースタイプの論理記憶領域から、前記制御部が管理する各論理記憶領域のインタフェースタイプに基づいて選択された前記第二のインタフェースタイプの論理記憶領域への、データのコピー処理を実行することを特徴とする方法。

10

【請求項 9】

請求項 6 又は 7 記載の方法であって、

前記制御部は、前記複数のディスク装置に対してコマンドを発行し、当該コマンドの応答として当該複数のディスク装置から当該複数のディスク装置のインタフェースタイプの情報を取得することを特徴とする方法。

【請求項 10】

複数のディスク装置と、

前記複数のディスク装置に接続される制御部とを有し、

前記複数のディスク装置には、各々が第一のタイプのインタフェースを有する複数の第一のインタフェースタイプのディスク装置と、各々が第二のタイプのインタフェースを有する複数の第二のインタフェースタイプのディスク装置とが含まれる記憶装置システムのも、前記制御部が有するプロセッサによって実行されるプログラムであって、

前記複数のディスク装置のインタフェースタイプに関する情報を取得する手順と、

前記複数のディスク装置のインタフェースタイプに関する情報を出力する手順と、

前記複数のディスク装置の中から出力された前記複数のディスク装置のインタフェースタイプに関する情報を用いてディスク装置で R A I D グループを作成するよう指示を受ける手順と、

20

30

前記複数の第一のインタフェースタイプのディスク装置から第一のインタフェースタイプの R A I D グループを、前記複数の第二のインタフェースタイプのディスク装置から第二のインタフェースタイプの R A I D グループを構成する前記指示に基づいて、前記選択されたディスク装置を用いて前記複数の第一のインタフェースタイプのディスク装置から第一のインタフェースタイプの R A I D グループを、前記複数の第二のインタフェースタイプのディスク装置から第二のインタフェースタイプの R A I D グループを構成する手順とを、前記プロセッサに実行させるためのプログラム。

【請求項 11】

請求項 10 記載のプログラムであって、

前記第一のタイプのインタフェースは F C インタフェース、前記第二のタイプのインタフェースは A T A インタフェースであることを特徴とするプログラム。

40

【請求項 12】

請求項 10 又は 11 記載のプログラムであって、

前記第一のインタフェースタイプの R A I D グループには第一のインタフェースタイプの論理記憶領域が、前記第二のインタフェースタイプの R A I D グループには第二のインタフェースタイプの論理記憶領域が構成されており、

前記プログラムは更に、

各論理記憶領域のインタフェースタイプを管理する手順と、

前記第一のインタフェースタイプの論理記憶領域から、前記制御部が管理する各論理記

50

憶領域のインタフェースタイプに基づいて選択された前記第二のインタフェースタイプの論理記憶領域への、データのコピーを行う手順とを、前記プロセッサに実行させることを特徴とするプログラム。

【請求項 13】

請求項 10 又は 11 記載のプログラムであって、

前記複数のディスク装置に対してコマンドを発行する手順と、前記コマンドの応答として前記複数のディスク装置から当該複数のディスク装置のインタフェースタイプに関する情報を取得する手順とを、前記プロセッサに実行させることを特徴とするプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は主として、記憶装置システムの制御方法及び記憶装置システムに関する。

【0002】

【従来の技術】

計算機に接続される記憶装置システムの種類に、ディスクアレイがある。ディスクアレイは、RAID (Redundant Arrays of Inexpensive Disks) と呼ばれ、アレイ状に配置された複数のディスク装置及びそれらを制御する制御部とを有する記憶装置システムである。ディスクアレイでは、リード要求（データの読み出し要求）およびライト要求（データの書き込み要求）がディスク装置の並列動作によって高速に処理され、かつデータに冗長性が付加される。ディスクアレイは、非特許文献 1 に開示されているように、付加される冗長データの種別とその構成により 5 つのレベルに分類されている。尚、ディスク装置とは、ハードディスクや光磁気ディスク等を有する記憶装置である。

【0003】

一方、ディスク装置には、他の機器との接続に使用するインターフェースが異なるいくつかの種類がある。例えば、ANSI T11 によって標準化作業が行われているファイバチャネル (Fibre Channel、以下「FC」) インターフェースを備えたディスク装置 (以下「FC ディスク装置」) や、ANSI T13 によって標準化作業が行われている ATA (ATA Attachment) インターフェースを備えたディスク装置 (以下「ATA ディスク装置」) である。

【0004】

ATA ディスク装置は、比較的安価で、家庭用のデスクトップコンピュータなどに利用される。一方、FC ディスク装置は、信頼性が要求される業務サーバなどで利用される。このため、FC ディスク装置は、ATA ディスク装置と比較して、データの入出力 (以下「I/O」) の処理性能が高く、かつ、24 時間 365 日の連続動作に耐える高い信頼性を有する。

【0005】

これらのディスク装置をディスクアレイ等の記憶装置システムに使用する場合、ディスクアレイに要求される性能、価格等に応じてディスク装置を使い分けることがある。例えば、データを格納するディスク装置よりも高性能なディスク装置を、冗長データであるパリティ格納用のディスク装置として用いることで、ディスクアレイに格納されるデータの信頼性の向上を図る技術が、特許文献 1 に開示されている。

【0006】

【特許文献 1】

特開平 10 - 301720 号公報

【非特許文献 1】

デビット・エー・パターソン (David A. Patterson) 外 2 名著、「レイド装置の一考察 (A Case for Redundant Arrays of Inexpensive Disks (RAID))」、1998 年 ACM SIGMOD 国際会議予稿集 (Proceedings of the 1988 ACM SIGMOD International Conference on Management of Data)、ACM プレス (ACM Press)、1988 年

10

20

30

40

50

【 0 0 0 7 】

【 発明が解決しようとする課題 】

近年、データ管理のコストを削減するために、複数のディスクアレイを物理的に一つに集約したいという要求がある。しかし、単にディスクアレイを物理的に一つに集約すると、種類や性能の異なる複数のディスクアレイを使用する場合に比べて融通性が損なわれる。そこで、高価で高信頼・高性能のディスク装置と、安価で信頼性の低いディスク装置とを一つのディスクアレイで混在させて用途ごとに使い分けることにより、ディスクアレイの価格を抑えつつ、より柔軟なアレイ構築を行いたいという要求が高まっている。

【 0 0 0 8 】

しかし、従来では、ディスク装置の信頼性を統一的に管理することが困難なことから、ディスク装置自体の信頼性、例えば装置寿命が異なるディスク装置を混在させてディスクアレイ等の記憶装置システムを構築し管理することは考えられていなかった。

10

【 0 0 0 9 】

本発明の目的は、上記課題を解決した記憶装置システム及び記憶装置システムの制御方法を提供することである。

【 0 0 1 0 】

【 課題を解決するための手段 】

本願発明は以下の構成を有する。すなわち、装置自体の信頼性が異なる複数の記憶装置及びこれら複数の記憶装置を制御する制御部を有する記憶装置システムである。ここで、制御部は、これら複数の記憶装置が接続されるインターフェース、記憶装置自体の信頼性を検出するための手段及び検出した情報を管理する手段を有する。例えば、検出手段としては、各記憶装置の種類を検出することによってその記憶装置の寿命（使用可能時間）を決定する手段が考えられる。管理手段としては、検出した記憶装置の寿命及びその記憶装置の使用時間を管理することで、ある時点でのそのディスク装置の使用可能時間（余命時間）を管理する手段が考えられる。

20

【 0 0 1 1 】

更に、本願発明の記憶装置システムには、管理手段によって管理されている記憶装置の余命時間に応じて、その記憶装置の使用形態を変更する手段が含まれても良い。例えば、管理されている記憶装置の残余時間がある所定の時間より短くなったら、その記憶装置をデータのバックアップ取得用として使用するようにする、あるいはその記憶装置の使用を停止する手段が考えられる。

30

【 0 0 1 2 】

又、本願発明の記憶装置システムには、管理手段によって管理されている記憶装置の使用可能時間に応じて、予めその記憶装置の用途を決定する手段が含まれていても良い。例えば、使用可能時間が長い記憶装置と短い記憶装置を区別し、使用可能時間が長い記憶装置群を通常使用する記憶装置とし、使用可能時間が短い記憶装置群をバックアップデータを格納する記憶装置とする手段が考えられる。

【 0 0 1 3 】

更に、本願発明には、検出手段によって検出された記憶装置の種類に基づいて、これらの記憶装置を制御する手段が含まれても良い。例えば、検出された記憶装置の種類に応じて、この記憶装置に特有のコマンドを発行する手段を設けても良い。

40

【 0 0 1 4 】

又、本願発明の記憶装置システムにおいて、記憶装置の種類を認識し、異なる記憶装置を別々に管理するプログラムであるドライブ管理手段とドライブ管理手段が用いる管理情報を格納したドライブ管理テーブルを設ける構成としても良い。

【 0 0 1 5 】

更に、本願発明の記憶装置システムは、累積時間の管理を行う以下の手段を有していても良い。例えば、ドライブの寿命を設定するドライブ寿命設定手段と、ATAディスク装置を計画的に起動/停止させるドライブ起動/停止手段と、動作の累積時間を計測する動作時間計測手段である。

50

【 0 0 1 6 】

又、上記のような記憶装置システムが混在する環境を管理するストレージ管理サーバを有する構成としても良い。

【 0 0 1 7 】

更に、記憶装置システム間又は記憶装置システム内におけるデータの複製を行う手段と、これらのデータ複製の対象となる論理記憶領域を寿命管理に基づいて割り当てる手段を有していても良い。

【 0 0 1 8 】

尚、「記憶装置」には、論理的な記憶装置及びディスク装置等の物理的な記憶装置の双方が含まれる。

10

【 0 0 1 9 】

【 発明の実施の形態 】

図 1 は、本発明を適用した記憶装置システムの第 1 の実施形態を示す図である。記憶装置システムの一例であるディスクアレイ 7 0 0 は、コンピュータ（以下「ホスト」）1 0 0 及び管理端末 5 0 0 と接続されている。ホスト 1 0 0 等との接続に使用されるインターフェースは、F C や SCS I が考えられる。ディスクアレイ 7 0 0 は、ディスクアレイ制御部 2 0 0、F C ディスク群 3 1 0 及び A T A ディスク群 4 1 0 とを有する。

【 0 0 2 0 】

F C ディスク群 3 1 0 は、複数の F C ディスク装置 3 0 1 を有する。A T A ディスク群 4 1 0 は、複数の A T A ディスク装置 4 0 1 を有する。尚、本実施形態では、F C ディスク装置及び A T A ディスク装置を例示するが、本発明の適用に当っては、ディスクアレイに採用されるディスク装置は、これらのディスク装置に限られない。

20

【 0 0 2 1 】

管理端末 5 0 0 は、ディスクアレイ 7 0 0 に関する設定を入力する際にユーザに使用されるキーボード装置等の入力部 5 1 0 及びディスクアレイ 7 0 0 に関する情報をユーザに示すディスプレイ装置等の出力部 5 2 0 とを有する。

【 0 0 2 2 】

ディスクアレイ制御部 2 0 0 は、ディスクアレイ 7 0 0 の制御をするためのプログラムを実行する C P U 2 1 0、メモリ 2 0 2、ホスト 1 0 0 からの入出力データを一時的に格納するキャッシュ 2 0 3、ホスト 1 0 0 とディスクアレイ制御部 2 0 0 との間のデータの送受信の制御を行うホスト F C I / F 2 0 4、F C ディスク群 3 1 0 とディスクアレイ制御部 2 0 0 との間のデータの送受信を制御するドライブ F C I / F 2 0 5、A T A ディスク群 4 1 0 とディスクアレイ制御部 2 0 0 との間のデータの送受信を制御するドライブ A T A I / F 2 0 6、管理端末 5 0 0 とディスクアレイ制御部 2 0 0 との間の情報の送受信を制御する管理 I / F 2 0 7 を有する。これらの部品は、内部バスで相互に結線される。

30

【 0 0 2 3 】

ドライブ F C I / F 2 0 5 は、F C ディスク群 3 1 0 と、F C を介して接続される。尚、本実施形態では、F C に用いるプロトコルとして、ファイバチャネルアービトラションループが用いられるが、これに限られず、他のプロトコル、例えば、F C のポイント・ツー・ポイント、ファブリック等のプロトコルが用いられても良い。

40

【 0 0 2 4 】

ドライブ A T A I / F 2 0 6 は、A T A ディスク群 4 1 0 に含まれる各 A T A ディスク装置と A T A バスを介して接続される。

【 0 0 2 5 】

メモリ 2 0 2 には、ディスクアレイ 7 0 0 を制御するために C P U 2 0 1 で実行される種々のプログラムが格納される。具体的には、ディスクアレイ 7 0 0 の動作を制御する R A I D 制御プログラム 2 1 0 及びディスクアレイ 7 0 0 の構成を管理する管理エージェント 8 0 0 である。又、メモリ 2 0 2 には、各種管理情報も格納される。具体的には、F C ディスク群 3 1 0 と A T A ディスク群 4 1 0 に関する情報が記録されるドライブ管理テーブ

50

ル240及びFCディスク群310に作成される論理的記憶領域(以下「FC LU」)320及びATAディスク群410に作成される論理的記憶領域(以下「ATA LU」)420に関する情報を記録するLU管理テーブル245である。

【0026】

RAID制御プログラム210は、更に幾つかのサブプログラムから構成される。具体的には、FCディスク群310とATAディスク群410にコマンドを発行する際にCPU201で実行されるドライブコマンド発行プログラム220、FCディスク群310とATAディスク群410の管理を行うために実行されるドライブ管理プログラム230、各ディスク群に設定された論理的記憶領域(以下「LU」)の管理を行うために実行されるLU管理プログラム225である。

10

【0027】

管理エージェント800は、更に幾つかのサブプログラムから構成される。具体的には、管理端末500からの入力を受けて、ディスク装置に関する情報(以下「ドライブ情報」)を設定する際に実行されるドライブ情報設定プログラム250、ドライブ情報を管理端末500へ出力する際に実行されるドライブ情報通知プログラム260、管理端末500からの入力を受けて、ディスクアレイ700に関する情報(以下「アレイ情報」)を設定する際に実行されるアレイ情報設定プログラム270及び管理端末500へアレイ情報を出力する際に実行されるアレイ情報通知プログラム280である。

【0028】

図2は、ドライブ管理テーブル240の内容の一例を示す図である。ドライブ管理テーブル240は、ディスクアレイ700が有するディスク装置毎に、種々の情報を登録する項目を有している。具体的には、ディスク装置に振られた番号が登録される「ドライブNo.」欄241、ディスク装置の種別が登録される「ドライブ種別」欄242、ディスク装置が含まれるRAIDグループ(冗長性を構成するディスク装置の組)及びそのRAIDレベル(以下「アレイ構成」)が登録される「アレイ構成」欄243、ディスク装置が含まれるRAIDグループのユーザ用途(以下「アレイ用途」)が登録される「用途」欄244、ディスク装置が起動しているか停止しているかの状況が登録される「起動状況」欄245、ディスク装置の累積動作時間が登録される「累積時間」欄246、ディスク装置の寿命が登録される「寿命設定」欄247、ディスク装置の余命が登録される「余命時間」欄248、並びに、ディスク装置が寿命に達したと判断されたときに自動的に他のディスク装置にデータを移行する(以下「自動移行」)かどうかの情報が登録される「自動移行指定」欄249がある。

20

30

【0029】

「ドライブ種別」欄242には、具体的には、FC、ATA等を示す情報が登録される。「アレイ構成」欄243には、RAIDグループを作成した順番を示す番号とRAIDレベルを示す情報が登録される。例えば、「(1)RAID5」が登録されたディスク装置は、1番目に作成されたRAIDレベル5のRAIDグループに含まれるディスク装置である。尚、ディスクアレイ700では複数のRAIDグループを作成すること、例えばRAID1とRAID5のRAIDグループを作成することが可能である。又、RAIDグループに含まれるディスク装置は、ディスクアレイ700が有するディスク装置全てでも一部でも良い。

40

【0030】

「用途」欄244には、具体的には、データベース用途を示すDB、ファイルシステム用途を示すFS等の情報が登録される。「起動状況」欄245には、ディスク装置が起動している場合はONが、ディスク装置が停止している場合はOFFを示す情報が登録される。「余命時間」欄248には、「寿命設定」欄247と「累積時間」欄246に登録された時間の差分を示す値が格納される。「自動移行指定」欄249には、当該ディスク装置に格納されたデータを自動移行するよう設定する場合は1、設定しない場合は0が登録される。

【0031】

50

ドライブ管理プログラム 230 は、ディスクアレイ制御部 200 が、ディスク装置の種類を識別する場合、ディスク装置の動作時間を計測する場合、ディスク装置の起動と停止を制御する場合、及びディスク装置が寿命に達したときの自動移行を実施する場合に CPU 201 で実行される。

【0032】

ドライブ情報設定プログラム 250 は、ディスクアレイ制御部 200 が、管理端末 500 からの入力に従って、ディスク装置の種類をドライブ管理テーブル 240 に設定する場合、ディスク装置の寿命をドライブ管理テーブル 240 に設定する場合、及びディスク装置の起動と停止の状況をドライブ管理テーブル 240 に設定する際に CPU 201 で実行される。

10

【0033】

ドライブ情報通知プログラム 260 は、ディスクアレイ制御部 200 が、管理端末 500 に対して、ディスク装置の種類を通知する場合、ディスク装置の累積動作時間を通知する場合、ディスク装置の起動と停止の状況を通知する場合、ディスク装置の交換を指示する場合、及びディスク装置が寿命に達したため他のディスク装置にデータを自動移行したことを通知する場合に、CPU 201 で実行される。

【0034】

アレイ情報設定プログラム 270 は、ディスクアレイ制御部 200 が、管理端末 500 からの入力に従い、アレイ構成を設定する場合及びアレイ用途を設定する場合に、CPU 201 で実行される。

20

【0035】

アレイ情報通知プログラム 280 は、ディスクアレイ制御部 200 が、管理端末 500 に対して、アレイ構成に関する情報を通知する場合及びアレイ用途に関する情報を通知する場合に、CPU 201 で実行される。

【0036】

図 3 は、ディスクアレイ 700 のユーザ又は管理者がディスクアレイ 700 の設定を行う際に、管理端末 500 の出力部 520 に表示される画面の一例を示した図である。この場合、出力部 520 には、以下の領域が表示される。すなわち、入力されたディスク装置の種別を表示するドライブ種設定領域 551、入力されたディスク装置の寿命を表示するドライブ寿命設定領域 552、ユーザによって指定されたディスク装置の起動/停止の条件を表示するドライブ起動/停止設定領域 553、入力されたアレイ構成を表示するアレイ構成設定領域 571 と、入力されたアレイ用途を表示するアレイ用途設定領域 572 である。ユーザ等は、この画面を見ながら、必要なデータを入力部を介して入力し、その内容を確認する。

30

【0037】

図 4 は、ディスクアレイ 700 のユーザ又は管理者が管理端末 500 を介してディスクアレイ 700 の情報を取得する際に、出力部 520 において表示される画面の一例を示したものである。この場合、出力部 520 には、以下の領域が表示される。

【0038】

すなわち、ディスクアレイ 700 が有するディスク装置の種別を表示するドライブ種表示領域 561、ディスクアレイ 700 が有するディスク装置の累積動作時間を表示する動作時間表示領域 562、ディスク装置の起動/停止の状況を表示するドライブ状態表示領域 563、ディスク装置の交換指示を表示する交換指示表示領域 564、あるディスク装置が寿命に達したため他のディスク装置にデータが自動移行されたことを表示する自動移行通知表示領域 565、設定されているアレイ構成を表示するアレイ構成表示領域 581 及び設定されているアレイ用途を表示するアレイ用途表示領域 582 である。管理者又はユーザは、この画面を管理端末 500 に表示させることで、ディスクアレイ 700 の状態を確認することができる。

40

【0039】

図 5 は、ディスクアレイ制御部 200 が、FC LU 320 及び ATA LU 420 を管理

50

するために使用するLU管理テーブル245の内容の一例を示す図である。LU管理テーブル245には、ディスクアレイ制御部200が各LUを管理するために、以下のような情報を登録する項目が各LU毎に存在する。

【0040】

「LU No.」欄2451には、各LUに対して任意に割り当てられた識別子が登録される。「ホスト割当状況」欄2452には、LUがホスト100に割り当てられているかどうか(以下割り当てられている場合には「使用済」、割り当てられていない場合には「未使用」という)を示す「有」又は「無」の情報が登録される。「SCSI ID」欄2453には、LUがホスト100に割り当てられている場合に、当該LUを含む論理的記憶装置に割り当てられるSCSI ID番号が登録される。「LUN」欄2454には、LUがホスト100に割り当てられている場合、ホスト100が当該LUをアクセスするために必要なSCSIロジカルユニット番号が登録される。

10

【0041】

「容量」欄2455には、LUに割り当てられている記憶容量が登録される。「LU種別」欄2456には、LUがFC LU320かATA LU420かという種別を示すための情報が登録される。「LU余命時間」欄2457には、LUの余命時間が登録される。具体的には、ディスクアレイ制御部200は、「ドライブNo.リスト」欄2458に登録された、LUを構成する各ディスク装置の余命時間をドライブ管理テーブル240から取得し、そのうちで最も小さい値を登録する。

【0042】

「ドライブNo.リスト」欄2458には、LUを構成するディスク装置を示す情報が、ドライブNo.の一覧として登録される。「評価値」欄2459には、LU選択の基準となる評価値を示す情報が登録される。評価値の計算方法等については、後述する。尚、使用済みLUの場合は、「評価値」欄2459には、当該欄に登録しうる最大値(本図の例では+99999)が登録される。

20

【0043】

以下、ディスクアレイ制御部200が、ユーザ又は管理者の指示に従って、ドライブ管理テーブル240、LU管理テーブル245及びRAIDグループの構築を行う場合について説明する。

【0044】

ディスクアレイ700に電源が投入された際に、ディスクアレイ制御部200は、ドライブ管理プログラム230を実行して、ドライブFC I/F205及びドライブATA I/F206に接続されているディスク装置の探査を実施する。ディスク装置の探査の一例としては、Mode Selectコマンドの使用する例が有る。具体的には、まず、ディスクアレイ制御部200がMode Selectコマンドを各ディスク装置に発行する。Mode Selectコマンドを受信したディスク装置は、当該ディスク装置に格納されたページ情報をディスクアレイ制御部200に送信する。ページ情報には、当該ディスク装置の種別に関する情報が含まれている。ディスク装置から取得したページ情報から、ディスクアレイ制御部200は、探査されたディスク装置の種類を識別する。

30

【0045】

更に、ディスクアレイ制御部200は、ドライブ情報設定プログラム250を実行して、検出されたディスク装置全てについて、ドライブ管理テーブル240の「ドライブNo.」及び「ドライブ種別」項目に、該当する情報を登録する。

40

【0046】

尚、ディスク装置の種類は、ユーザが管理端末500の入力部を用いて入力しても良い。この場合、ディスクアレイ制御部200は、管理端末500から受取った情報に基づいて、ドライブ管理テーブル240に情報を登録する。

【0047】

ディスク装置の種別を検出又は登録した後、ディスクアレイ制御部200は、ユーザの指示に基づいて、アレイ構成やアレイ用途の設定を行う。

50

【 0 0 4 8 】

まず、ユーザが管理端末500の入力部510を用いて入力した「ドライブNo.」「ドライブ種別」の情報取得の指示を受信すると、ディスクアレイ制御部200は、ドライブ情報通知プログラム260を実行して、要求された情報をドライブ管理テーブル240から取得し、管理I/F207を介して管理端末500に転送する。管理端末500は、受信した情報を、出力部520に表示されるドライブ種表示領域561の「ドライブNo.」及び「ドライブ種別」を表示する部分に表示する。尚、ユーザがディスク装置の種別を定義する場合には、本処理は省略されて構わない。

【 0 0 4 9 】

その後、ユーザは、管理端末500の出力部520に表示されたドライブ種表示領域561に表示された情報に基づいて任意のディスク装置を選定し、アレイ構成設定画面571を用いて、選定されたディスク装置を用いてRAIDグループを構築する指示を入力部510から入力する。また、ユーザは、構築するRAIDグループの用途を入力する。ディスクアレイ制御部200は、アレイ情報設定プログラムを実行して、管理I/F207を介して受信したRAIDグループや用途等のアレイ情報を、ドライブ管理テーブル240の「アレイ構成」「用途」欄に登録する。

10

【 0 0 5 0 】

ユーザは、ドライブ種表示画面561に表示されたディスク装置の種別にもとづきRAIDグループを構築するが、この際、信頼性、性能の低いATAディスク装置を、高信頼、高性能が要求されるアレイ用途、例えばデータベース向けのアレイ用途に設定しないようにすることが可能である。具体的には、後述するディスク装置の寿命の情報に基づいて、ユーザから「用途」が指定された場合に、その用途にディスク装置の寿命が合致するかどうかを判断する。そして、その判断結果に従って、ユーザに警告を行う。

20

【 0 0 5 1 】

例えば、ATAディスク装置を高信頼、高性能なアレイ用途に設定しようとした場合に、管理端末500が、出力部520を介してユーザに注意、警告を発行するようにしてもよい。また、ドライブ種表示画面561にてディスク装置を表示する場合、種類が異なるディスク装置を区別できるように別々の図柄にしてもよい。

【 0 0 5 2 】

また、最初にRAIDグループを作成する際には、ディスク装置の寿命については考慮せず、再度RAIDグループを作り直す際に、ディスク装置の余命を新たな用途設定の判断基準として加えるという方法も考えられる。

30

【 0 0 5 3 】

尚、ユーザが管理端末500の入力部510より入力した「アレイ構成」「用途」の情報取得を指示する情報を受取ると、ディスクアレイ制御部200は、アレイ情報通知プログラムを実行することで、管理I/F207を介して管理端末500から要求された情報を、ドライブ管理テーブル240から検索し、検索された情報を転送する。情報を受取った管理端末500は、出力部520のアレイ構成表示領域581及びアレイ用途表示領域582に取得した情報を表示する。

【 0 0 5 4 】

次に、ディスクアレイ制御部200が行うディスク装置の寿命管理について説明する。前述したように、ディスクアレイ制御部200は、ドライブ管理プログラム230を実行することで、ディスク装置の動作累積時間の管理、計画的な起動/停止及びデータの自動移行を制御する。

40

【 0 0 5 5 】

ここで、ディスク装置の「寿命」とは、ディスク装置の平均故障間隔や推奨稼働時間などの仕様を元にディスクアレイの製造元が設定した時間、または、ディスクアレイ製造元が設定した時間を元にディスクアレイ管理者が設定した時間を指す。ここでは、「ディスク装置の長時間稼働が原因となる故障を防ぐための、予防保守の基準となる時間」と定義する。したがって、「寿命」を上回ったディスク装置については、格納したデータの移行を

50

実施した上で、保守員等により新たなディスク装置への交換を実施する必要がある。

【 0 0 5 6 】

寿命管理を行うために、まず、ディスクアレイ 7 0 0 は、ディスク装置ごとの寿命設定や自動移行指定を行う。本処理は、ディスクアレイ制御部 2 0 0 が、各種テーブルの設定をユーザの指示に従って行う際に一緒に行われてもよい。あるいは、その前に行われてもよい。ディスク装置ごとの寿命設定、自動移行指定は以下のように行われる。

【 0 0 5 7 】

ユーザが管理端末 5 0 0 の入力部 5 1 0 から入力した「ドライブ No.」「ドライブ種別」の情報取得の指示を受信すると、ディスクアレイ制御部 2 0 0 は、管理 I / F 2 0 7 を介して管理端末 5 0 0 に要求された情報を転送する。情報を受信した管理端末 5 0 0 は、出力部 5 2 0 のドライブ種表示領域 5 6 1 に、受信した「ドライブ No.」及び「ドライブ種別」についての情報を表示する。

10

【 0 0 5 8 】

ユーザは、ドライブ種表示領域 5 6 1 に表示された情報に基づいてディスク装置の種類を識別し、管理端末 5 0 0 の入力部 5 1 0 を介して、ディスク装置の寿命及び自動移行の要否を設定する。管理 I / F 2 0 7 を介してディスク装置の寿命及びディスク装置の自動移行の要否に関する情報を受信したディスクアレイ制御部 2 0 0 は、ドライブ情報設定プログラム 2 5 2 を実行して、ドライブ管理テーブル 2 4 0 の「寿命設定」及び「自動移行指定」の欄に、受信した情報を登録する。なお、「寿命設定」及び「自動移行指定」の情報は、ユーザが指定せずに、ディスクアレイ制御部 2 0 0 がディスク装置の種類より一意に決定して、ドライブ管理テーブル 2 4 0 に設定してもよい。例えば、FC ディスク装置だったら最初から寿命が長いと予測して、ドライブ管理テーブル 2 4 0 に設定することが考えられる。

20

【 0 0 5 9 】

図 6 は、ディスクアレイ制御部 2 0 0 が行うドライブ寿命管理の手順を示すフローチャートである。

【 0 0 6 0 】

まず、ディスクアレイ制御部 2 0 0 は、ドライブ管理プログラム 2 3 0 を実行することで、ディスクアレイ 7 0 0 が有する各ディスク装置の動作時間の計測を開始し、結果をドライブ管理テーブル 2 4 0 の「累積時間」に積算していく。尚、動作時間を計測する場合、ディスクアレイ 7 0 0 が有するクロックに基づいて動作時間の計測を行う（ステップ 1 0 0 1）。その後、ディスクアレイ制御部 2 0 0 は、ドライブ管理テーブル 2 4 0 の「累積時間」と「寿命設定」との値を比較し、累積時間が設定された寿命に到達したかどうかを判定する。尚、比較のタイミングは、他の処理とは非同期に任意のタイミングで構わない（ステップ 1 0 0 4）。

30

【 0 0 6 1 】

累積時間が寿命に到達していない場合は、ステップ 1 0 0 1 の処理に戻る。累積時間が設定された寿命に到達した場合は、ディスク装置の「自動移行指定」欄 2 4 9 に 1 が登録されているかどうかを判定する（ステップ 1 0 0 5）。自動移行の指定がされていない場合（登録された値が 0）、ディスクアレイ制御部 2 0 0 は、ドライブ情報通知プログラム 2 6 0 を実行することで、累積時間が設定された寿命に到達したディスク装置を、交換すべきディスク装置として、管理 I / F 2 0 7 を介して管理端末 5 0 0 に通知する。

40

【 0 0 6 2 】

管理端末 5 0 0 は、出力部 5 2 0 の交換指示表示領域 5 6 4 に、累積時間が寿命に達したディスク装置を交換すべきディスク装置として表示する（ステップ 1 0 0 6）。累積時間が寿命に達したディスク装置について自動移行の指定がある場合（登録された値が 1）、ディスクアレイ制御部 2 0 0 は、ドライブ管理プログラム 2 3 0 及び R A I D 制御プログラム 2 1 0 を実行することで、累積時間が寿命に達したディスク装置に格納されているデータを、自動移行先となるディスク装置に転送する。尚、自動移行先となるディスク装置は、ユーザ等によりあらかじめ指定されている。この場合、1 つのディスク装置が、複数

50

のディスク装置の自動移行先ディスク装置として指定されていても良い(ステップ1010)。

【0063】

データの転送が完了した後、ディスクアレイ制御部200は、管理I/F270を介して、管理端末500に自動移行先となったディスク装置に関する情報を通知する(ステップ1011)。情報が通知されると、管理端末500は、出力部520の自動移行通知表示領域565に、自動移行先となったディスク装置に関する情報を表示する。ドライブ管理手段230は、自動移行したドライブに関するドライブ管理テーブル240の「累積時間」をクリアし(ステップ1012)、ステップ1003に戻る。

【0064】

尚、本実施形態では累積時間が寿命に達したディスク装置について自動移行等を行うが、余命時間がある閾値を下回ったディスク装置について自動移行等を行っても良い。この場合、ディスクアレイ制御部200は、「余命時間」2408に登録された値を確認して、閾値を下回るディスク装置を特定する。

【0065】

一方、ディスクアレイ制御部200は、ディスク装置の累積動作時間の増加を抑止するため、ディスク装置の動作の起動/停止を制御する。この制御は、図11で示した寿命管理とは独立に行われる。

【0066】

ディスクアレイ制御部200は、予め設定された契機に基づいて、ディスク装置の起動/停止を行う。ディスク装置を停止/起動させる契機は、たとえば、夜間等のユーザからのアクセスが完全に止められる時間とすることができる。あるいは、あるディスク装置の余命時間がある閾値より短くなった場合をディスク装置の起動/停止の契機とすることもできる。

【0067】

また、ユーザが、ディスク装置の起動/停止を管理端末500の入力部510を介して直接指定してもよい。管理端末500は、ディスク装置の起動/停止に関する指示を、管理I/F207を介して、ディスクアレイ制御部200に送る。ディスクアレイ制御部200は、受信した指示に基づいて、ディスク装置の起動/停止を制御する。その後、ディスクアレイ制御部200は、ディスク装置の起動/停止状況をドライブ管理テーブル240の「起動状況」に登録する。なお、ドライブの起動/停止は、ドライブにStart/StopUnitコマンドを発行することで行なう。

【0068】

尚、ユーザは、ドライブの起動状況、累積動作時間を取得する場合、管理端末500の入力部510より「起動状況」「累積時間」の情報取得を指示する。ディスクアレイ制御部200は、管理I/F207を介して管理端末500に指定された情報を転送する。管理端末500は、出力部520のドライブ状態表示領域563及び動作時間表示領域562に、受信したディスク装置の起動状況及びディスク装置の累積動作時間を表示する。

【0069】

ユーザは、表示されたドライブの累積動作時間、または、余命時間にもとづき、あらかじめディスク装置の交換を実施してもよい。

【0070】

次に、異なるインターフェースを有するディスク装置に対してディスクアレイ制御部200がコマンドを発行する場合について説明する。

【0071】

例えば、ATAディスク装置401は、複数のタグをつけたコマンドを処理することができない。そこで、ディスクアレイ制御部200は、ホスト100からデータの書き込み等を指示するコマンドを受信した際には、ドライブコマンド発行プログラム220を実行して、ディスク装置に発行すべきコマンドを決定する。具体的には、ディスクアレイ制御部200は、ディスク装置へコマンドを発行する際にドライブ管理テーブル230の「ドラ

10

20

30

40

50

イブ種別」を参照して、コマンドの発行先がFCディスク装置301なのかATAディスク装置401なのかを認識し、ATAディスク装置の場合、複数のタグをつけたコマンドを発行せず、単一のコマンドを発行する。

【0072】

本実施形態によれば、ディスク装置自身の信頼性、例えば、寿命が異なるディスク装置を混在させて、同一のディスクアレイ700内で管理することが可能となる。また、RAIDグループを構成するディスク装置の信頼性を考慮して、LUの用途を決定することができる。また、制御方式の違うディスク装置を混在させて、同一のディスクアレイ内で管理することが可能となる。

【0073】

図7は、本発明を適用した計算機システムの第二の実施形態を示す図である。以下、第一の実施形態と相違する点のみ述べる。図7において、ディスクアレイ700'は、FCディスク群310を格納するためのFCドライブ筐体300及びATAディスク群を格納するためのATAドライブ筐体400を有している。FCドライブ筐体300内に格納されたFCディスク装置301は、FCドライブ筐体300が有するポートバイパス回路330を介して、ドライブFC I/F205及びATAドライブ筐体400とに接続される。また、FCドライブ筐体300とATAドライブ筐体400との間は、FCループ600、具体的には、ファイバーケーブルで接続される。

【0074】

ATAドライブ筐体400に格納されたATAディスク装置401は、直接FCループ600と接続することができない。そこで、ATAドライブ筐体400は、FCインターフェースとATAインターフェースとを変換するためのFC-ATA変換 I/F430を有する。したがって、ATAディスク装置401は、FC-ATA変換 I/F430及びFCループ600を介して、FCドライブ筐体300と接続されることになる。

【0075】

図8は、ポートバイパス回路330の構成を示す図である。本図のポートバイパス回路330は5つのポートを有している。また、ポートバイパス回路330は、ポート数と同じセレクタ331を有し、これにより、ポートごとのループ結線を実施する。セレクタ331により、ポートに結線を接続した状態であればポートを経由した信号パスを構成し、ポートが未接続であればポートを経由しない内部回路を通した信号パスを形成する。

【0076】

図9は、FC-ATA変換 I/F430の構成を示す図である。FC-ATA変換 I/F430は、FC I/FにATAディスク装置を接続するためのものである。FC-ATA変換 I/F430は、FCとのインターフェースであるFC I/F431、ATAとのインターフェースであるATA I/F438、メモリ441、メモリ441内にあるプログラムを実行するCPU440、及びFC I/F431、ATA I/F438、メモリ441、及びCPU440を結線する内部バス442を有する。

【0077】

メモリ441には、FCのコマンドの受信の際に実行されるFCコマンド受信プログラム432、FCとのデータ送受信の際に実行されるFCデータ転送プログラム433、FCのコマンドをATA向けのコマンドに変換する際に実行されるコマンド変換プログラム434、FCとATAとの間のデータのバッファリングを行うデータバッファ435、変換したコマンドをATA側に発行する際に実行されるATAコマンド発行プログラム436、ATAとのデータ送受信の際に実行されるATAデータ転送プログラム437、及びコマンドを保持、再生成のためのテンポラリバッファリングを行なうコマンドバッファ439が格納される。

【0078】

CPU440は、FCコマンド受信プログラム432を実行してFCからのコマンドを受信し、FCデータ転送プログラム433を実行して、FC I/F431とデータバッファ435間のデータ転送を行なう。また、CPU440は、コマンド変換プログラム43

10

20

30

40

50

4 を実行して F C のコマンドを A T A 向けのコマンドに変換し、A T A コマンド発行プログラム 4 3 6 を実行して、変換したコマンドを A T A 側に発行する。また、C P U 2 2 0 は、A T A データ転送プログラム 4 3 7 を実行して、A T A I / F 4 3 8 とデータバッファ 4 3 5 間のデータ転送を行なう。

【 0 0 7 9 】

本実施形態では、F C - A T A 変換 I / F 4 3 0 を用いることで、F C を介して、ディスクアレイ制御部 2 0 0 と A T A ディスク装置 4 0 1 とを接続することができる。F C - A T A 変換 I / F 4 3 0 は、ディスクアレイ制御部 2 0 0 ' から受信したコマンドを A T A ディスク装置用のコマンドに変換する。また、ディスクアレイ制御部 2 0 0 ' と A T A ディスク装置との間で送受信されるデータも、F C - A T A 変換 I / F 4 3 0 によって変換される。

10

【 0 0 8 0 】

本実施形態によれば、F C ディスク装置を接続するインターフェースであるファイバチャネルアービトレートッドループに、A T A ディスク装置を接続することができる。

【 0 0 8 1 】

図 1 0 は、本発明を適用した計算機システムの第三の実施形態を示す図である。以下、第一の実施形態と異なる点を述べる。本実施形態の計算機システムは、第一の実施形態の構成に加え、ディスクアレイ 7 0 1 A に接続される他のディスクアレイ 7 0 1 B、ディスクアレイ 7 0 1 B に接続されるホスト計算機 1 0 0、ディスクアレイ 7 0 1 A 及び 7 0 1 B と管理用ネットワーク 2 0 0 0 0 を介して接続される計算機であるストレージ管理サーバ 1 0 0 0 0 を有する。2 台のディスクアレイは、各ディスクアレイが有するホスト F C I / F 2 0 4 及びリモートコピー接続 3 0 0 0 0 を介して相互に接続される。なお、リモートコピー接続 3 0 0 0 0 の方法には、例えばファイバチャネルや E S C O N などが考えられるが、本実施形態では特に区別しない。また、各ディスクアレイ 7 0 1 と管理用ネットワーク 2 0 0 0 0 との間の接続には、各ディスクアレイ 7 0 1 が有する管理 I / F 2 0 7 が用いられる。

20

【 0 0 8 2 】

また、本実施形態のディスクアレイ 7 0 1 は、第一の実施形態の構成に加え、新たに、ディスクアレイ制御部 2 0 0 1 内のメモリ 2 0 2 内に L U ペア管理テーブル 2 4 6 を有する。又、R A I D 制御プログラムに、ディスクアレイ 7 0 1 A と接続されるディスクアレイ 7 0 1 B との間のデータの複製（以下、「リモートコピー」）を制御するためのプログラムモジュール、及びディスクアレイ 7 0 1 A 又は B 内におけるデータの複製（以下「装置内レプリケーション」）を管理するためのプログラムモジュールが新たに追加されている。尚、リモートコピーや装置内レプリケーションの実施方法については、同期（データ複製が完了するまで上位装置にデータの複製完了を報告しない）、非同期（データ複製の完了を待たずに、上位装置にデータ複製の完了を報告する）といったバリエーションがあるが、本実施形態においては、特に区別しない。

30

【 0 0 8 3 】

図 1 1 は、L U ペア管理テーブル 2 4 6 の一例を示す図である。ディスクアレイ 7 0 1 A 及び 7 0 1 B は、それぞれのメモリ 2 0 2 に登録された L U ペア管理テーブル 2 4 6 を用いて、リモートコピーに使用されるディスクアレイ間で同じデータを保持する L U の対又は装置内レプリケーションで生成されるディスクアレイ 7 0 1 内にある同じデータを保持する L U の対（以下「L U ペア」）を管理する。

40

【 0 0 8 4 】

L U ペア管理テーブル 2 4 6 は、「L U ペア N o .」、「コピー元 L U 情報」、「コピー先 L U 情報」、及び「ペア状態」の項目を有する。

【 0 0 8 5 】

「L U ペア N o .」欄 2 4 6 a には、L U ペアに対して任意に割り当てられた識別子が登録される。本実施形態では、ディスクアレイ 7 0 1 間及びディスクアレイ 7 0 1 内の L U ペアを区別しないで識別子が割り当てられる。

50

【0086】

「コピー元LU情報」項目は、以下に説明する3つの下位項目を有する。「LU No.」欄246bには、ある識別子が与えられたLUペアのうち、オリジナルのデータが格納されたLU（以下「コピー元LU」）に割り振られたLU No.が登録される。「ペア属性」欄246cには、LUペアがリモートコピーに基づくLUペアであるのか、装置内レプリケーションに基づくLUペアかを識別するための情報、具体的には、リモートコピーの場合には「リモート」、装置内レプリケーションの場合には「ローカル」を示す値が登録される。「LU種別」欄246dには、コピー元LUを格納するディスク装置の種別が登録される。

【0087】

「コピー先LU情報」項目は、以下に説明する3つの下位項目を有する。「装置ID」欄246eには、コピー元LUと対になるLU（以下「コピー先LU」）が属するディスクアレイ701に割り振られた識別子が登録される。例えば、リモートコピーの場合にはリモートコピー先のディスクアレイ701に割り振られた識別子が、装置内レプリケーションの場合にはコピー元LUを有するディスクアレイ701に割り振られた識別子が登録される。「LU No.」欄246fには、コピー先LUに割り当てられたLU No.が登録される。「LU種別」欄246gには、コピー先LUを保持するディスク装置の種別が登録される。

【0088】

「ペア状態」欄246hには、LUペアの現在の状態が登録される。例えば、LUペアの各々のLUに格納されているデータの同期が取られ、各々に格納されたデータの内容が一致している状態（以下「Pair状態」）、又は、LUペア間でデータの同期が取られていない状態（以下「Split状態」）を示す値が登録される。

【0089】

ディスクアレイ701は、装置内レプリケーションの機能を用いて、以下のことを行う。例えば、Pair状態にあったLUペアを、任意の時間においてSplit状態に変更する。このようにすると、コピー先LUには、任意の時間におけるLUペアが有するデータが保存されることになる（このような処理を「スナップショットを取る」という）。その後、ホスト100がコピー先LUからデータを読み出して他の記憶装置、例えばテープ装置等に書込むことで、スナップショットを取得した時間におけるLUペアに格納されたデータのバックアップをとることができる。また、スナップショットを取得した後のコピー先LUそのものをデータのバックアップとして保存しておいてもよい。

【0090】

尚、リモートコピーのLUペアと装置内レプリケーションのLUペアを組み合わせることも可能である。例えば、ディスクアレイ701A及びBとの間でリモートコピーを行う場合に、ディスクアレイ701Bが有するコピー先LUを、ディスクアレイ701Bが実行する装置内レプリケーションのコピー元LUとする。そして、ディスクアレイ701Bの内部で、装置内レプリケーションのLUペアとなるコピー先LUを作成してもよい。ここで作成されたコピー先LUは、上述したように、split状態にすることで、データのバックアップに使用することができる。

【0091】

ユーザは、FC LU 320及びATA・LU 420を自由に組み合わせてコピー元LU及びコピー先LU（以下「LUペア」）を設定することが出来る。特に、ユーザは、後述するストレージ管理サーバ10000を介して、各LUの特徴を考慮したLUペアの割当を設定することができる。例えば、LU種別を特徴とし、コピー元LUに高性能、高信頼なFCディスク装置を使ったファイバチャネルLU 320、コピー先LUに低価格なATAディスク装置を使ったATA・LU 420を割り振ることができる。また、LU余命時間を特徴とし、余命時間が少ないディスク装置を有するLUは、割当の対象としないとか、装置内レプリケーションのコピー先にのみ使用するなども設定することができる。

【0092】

次に、ストレージ管理サーバ10000について説明する。図12は、ストレージ管理サーバ10000の構成を示す図である。

【0093】

ストレージ管理サーバ10000は、CPU11000、主記憶12000、記憶部13000、管理用ネットワーク20000と接続するための管理I/F14000、出力部15000及び入力部16000を有する。これらの部品は、バス等の通信路17000を介して相互に接続される。記憶部13000には、CPU11000で実行されるプログラムであるストレージマネージャ13100、ディスクアレイ701から収集した情報を格納する領域であるストレージリポジトリ13200、ストレージマネージャ13100が管理対象としているディスクアレイ700の一覧を保持する装置リスト13300、LU評価項目テーブル13400及びLUの余命管理を行うためのLU余命閾値テーブル13500とが格納されている。

10

【0094】

なお、ストレージリポジトリ13200には、ディスクアレイ701A及びBが有するドライブ管理テーブル240、LU管理テーブル245、及びLUペア管理テーブル246に格納された情報が定期的に収集され、これらのテーブルの複製が作成される。また、装置リスト13300、LU評価項目テーブル13400、及びLU余命閾値テーブル13500は、ユーザの指示に基づいてストレージ管理サーバ10000がストレージマネージャ13100を実行することにより作成される。

【0095】

図13は、装置リスト13300の一例を示す図である。装置リスト13300は、ストレージ管理サーバ10000が、自身が管理すべきディスクアレイ701を確認し、ディスクアレイ701と通信するために必要な情報を参照するために使用される。「対象装置No.」欄13300aには、ストレージ管理サーバ10000が管理するディスクアレイ701に対して任意に割り当てられた識別子が登録される。「装置ID」欄13300bには、ディスクアレイ701に割り振られた固有の識別子が登録される。「IPアドレス」欄13300cには、ストレージ管理サーバ10000がディスクアレイ701と通信するために必要な情報、例えば、管理用ネットワーク20000にIPネットワークが使用される場合には、ディスクアレイ700に割り振られたIPアドレスに関する情報が登録される。

20

30

【0096】

図14は、LU評価項目テーブル13400の一例を示す図である。LU評価項目テーブル13400には、LUペアのLUを選定する際に考慮されるべき項目とその項目における評価内容が登録されている。ストレージ管理サーバ10000は、ユーザの指示に合致するLUを検索する際に、LU評価項目テーブル13400に登録された項目で各LUを評価し、そのテーブルで指定された評価内容を、LU管理テーブル245の複製(以下、「LU管理テーブル245'」)の評価値2459の項目に登録する。その後、その評価結果に基づいて、ストレージ管理サーバ10000は、適当なLUを選択し、ユーザに提示する。

【0097】

「評価項目」欄13400aには、ユーザが選択できるLU指定時の評価項目が登録される。「評価対象LU管理テーブル欄」欄13400bには、評価項目で評価の対象となるLU管理テーブル245内の項目、例えば、LU種別2456が登録される。「入力条件読み替え」欄13400cには、LU管理テーブル245'に登録された各LUの評価を行うとき、ユーザが入力した条件の読み替えが必要か否かが登録される。「条件成立時ペナルティ」欄13400dには、ユーザが入力した条件と評価される対象のLU管理テーブル245'の項目に登録された内容が合致するとき、LU管理テーブル245'の「評価値」欄2459の値に加算される値が格納される。「条件不成立時ペナルティ」欄13400eには、ユーザが入力した条件と評価される対象のLU管理テーブル245の欄に登録された内容が合致しないときに、LU管理テーブル245'の「評価値」欄2459

40

50

の値に加算される値が登録される。尚、具体的なLUの評価方法については、後述する。

【0098】

図19は、LU余命閾値テーブル13300の一例を示す図である。LU余命閾値テーブル13300は、ストレージ管理サーバ10000が各ディスクアレイ701が有するLUの余命を管理する際に使用される。具体的には、以下の情報を登録する項目を有する。

【0099】

「閾値No.」欄13500aには、複数設定された閾値に対して各々任意に割り当てられた識別子が登録される。「装置ID」欄13500bには、余命管理の対象となるディスクアレイ701に割り当てられた識別子が登録される。「対象LUリスト」欄13500cには、管理対象となったディスクアレイ701が有するLUのうち、実際に余命が管理されるLUを示す情報が登録される。「余命閾値」欄13500dには、識別子に対応する閾値を示す情報が登録される。

10

【0100】

本実施形態においては、管理対象として登録されたLUについて、ストレージ管理サーバ10000が余命時間を確認し、確認した余命時間が設定された閾値を下回ると、ストレージ管理サーバ10000からユーザに警告が出力される。

場合によっては、ユーザへLU移行設定を促したり、ディスクアレイ装置に自動移行指示を発行したりする。尚、LUに格納されたデータの重要性が低い場合には、余命時間を判断基準にするのではなく、LUを構成するディスク装置の寿命時間を判断基準としても良い。ただし、この場合、ディスク装置の早急な交換が必要になる。

20

【0101】

本実施形態においては、ユーザは、ストレージ管理サーバ10000を用いて、入出力I/Fが異なるディスク装置を混在させたディスクアレイ701のLUを作成するにあたり、ディスク装置の性能や信頼性などのパラメータを考慮することができる。以下、この処理をLU割当処理と称する。なお、特に断りのない限り、以下のフローチャートで説明される処理は、ストレージ管理サーバ10000がストレージマネージャ13100を実行することで実施する処理である。

【0102】

図15は、LU割当処理2000の手順を示すフローチャートである。最初に、ストレージ管理サーバ10000は、ユーザからの指示により、コピー元となるLUをホスト1000に割り当てる処理か、リモートコピーや装置内レプリケーションにおけるコピー先LUを割り当てる処理を実行するかを判断する(ステップ2100)。

30

【0103】

ステップ2100でコピー元LUの割当と判断した場合、ストレージ管理サーバ10000は、ホスト提供LU割当処理を行う。本処理の詳細は後述する(ステップ2200)。その後、ストレージ管理サーバ10000は、コピー先LUの割当を行うかどうかをユーザに問い合わせる。具体的には、出力部15000に処理を行うかどうかを問い合わせる表示を出力する(ステップ2300)。

【0104】

ステップ2100でコピー先LU割当がユーザより指示された場合及びステップ2300でコピー先LU割当がユーザより指示された場合は、ストレージ管理サーバ10000は、コピー先LU割当処理を行う。本処理についても詳細は後述する(ステップ2400)。その後、ストレージ管理サーバ10000は、LU割当処理を終了する。ステップ2300でコピー先LU割当処理が不要と指示された場合は、ストレージ管理サーバ10000は、ただちにLU割当処理を終了する。

40

【0105】

図16は、図15のステップ2200にて実行されるホスト提供LU割当処理の詳細手順を示すフローチャートである。ストレージ管理サーバ10000は、ユーザから、コピー元LUに必要とされる条件(以下、「パラメータ」)を受け取る。ここでパラメータとは、LUが作成される対象装置、LUの記憶容量、ホストFC I/F、SCSI ID、L

50

UN等、LU作成に必要な基本情報に加え、LU評価項目テーブル13400の「項目」欄に登録された評価項目についての情報である(ステップ2210)。

【0106】

その後、ストレージ管理サーバ10000は、ユーザより与えられたパラメータをもとに、適当なLUをコピー元LUの候補として選択する割当候補LU探索処理を行う。本処理の詳細は後述する(ステップ2220)。その後、選択されたLUの候補をユーザに示し、ユーザの確認を求める(ステップ2230)。

【0107】

ステップ2230で、提示されたLU候補をユーザがコピー元LUとして承認した場合、ストレージ管理サーバ10000は、承認されたLUを有するディスクアレイ701に格納されたLU管理テーブル245及びその複製である245'を更新する。具体的には、最初に、ストレージ管理サーバ10000は、ストレージリポジトリ13200に格納されたLU管理テーブル245'について、承認されたLUの「ホスト割当状況」欄245bを「有」に変更し、「SCSI ID」欄245cおよび「LUN」欄245dに、ユーザから与えられたパラメータの値を登録する。その後、ストレージ管理サーバ10000は、LU管理テーブル245'の更新内容をLU管理テーブル245に反映するよう、承認されたLUを有するディスクアレイ701に指示をする(ステップ2240)。

【0108】

その後、ストレージ管理サーバ10000は、ディスクアレイ制御部200に対して、更新が反映されたLU管理テーブル245をもとにLU管理を実行することを指示し、ホスト提供LU割当処理を終了する(ステップ2250)。

【0109】

ステップ2230において、提示されたLUをユーザが却下した場合、ストレージ管理サーバ10000は、ホスト提供LU割当処理を中止する(ステップ2260)。尚、処理の自動化を図るために、選択されたLUをユーザに提示することなく、コピー元LUとして設定する処理が行われても良い。

【0110】

図17は、図15のステップ2400に示されるコピー先LU割当処理の詳細手順を示すフローチャートである。

【0111】

ストレージ管理サーバ10000は、ユーザからコピー先LUを選択するためのパラメータを受け取る。ここでパラメータとは、上述したコピー元LUで説明した情報である(ステップ2410)。

【0112】

その後、ストレージ管理サーバ10000は、ユーザより与えられたパラメータをもとに、コピー先LUとなりうる候補を選択する割当候補LU探索処理を行う。本処理の詳細は後述する(ステップ2420)。その後、ストレージ管理サーバ10000は、ステップ2420で選択されたコピー先LUの候補をユーザに示し、選択されたLUの使用の確認を要求する(ステップ2430)。

【0113】

ステップ2430において、提示されたLUをユーザがコピー先LUとして承認した場合、ストレージ管理サーバ10000は、LUペア管理テーブル246及びLU管理テーブル245と、それぞれの複製246'、245'を更新する。また、ストレージ管理サーバ10000は、更新した内容を各テーブルに反映するよう、選択されたLUを格納するディスクアレイ701に指示する。具体的には、LUペア管理テーブル246'については、選択されたコピー先LUについて新たにLUペアエントリを作成し、ユーザが指定したパラメータ及び選択されたコピー先LUの情報を登録する。またLU管理テーブル245'については、選択されたLUの「ホスト割当状況」欄245bを「有」に変更し、「SCSI ID」欄245cおよび「LUN」欄245dにパラメータの値を登録する(ステップ2440)。

10

20

30

40

50

【 0 1 1 4 】

その後、ストレージ管理サーバ10000は、選択されたコピー先LUとペアになるコピー元LUを有するディスクアレイ701のディスクアレイ制御部200に対して、更新後のペアLU管理テーブル246に基づいてリモートコピー処理あるいは装置内レプリケーション処理を実行することを指示し、コピー先LU割当処理を終了する(ステップ2450)。

【 0 1 1 5 】

ステップ2430において、提示されたコピー先LUをを却下した場合、ストレージ管理サーバ10000は、ホスト提供LU割当処理を中止し、ホスト提供LU割当処理を終了する(ステップ2460)。尚、処理の自動化を図るために、選択されたLUをユーザに提示することなく、コピー先LUとして設定する処理が行われても良い。

10

【 0 1 1 6 】

図18は、図16のステップ2220及び図17のステップ2420で示される、割当候補LU探索処理の詳細手順を示すフローチャートである。

【 0 1 1 7 】

最初に、ストレージ管理サーバ10000は、ストレージリポジトリ13200に格納されたLU管理テーブル245'の中から、ホスト100に使用されていないLUを抽出する。具体的には、LU管理テーブル245'の「ホスト割当状況」欄245bに「有」が登録されているLUは使用済みと判断され、評価値に最大値(図5では+9999)が登録される。「無」が登録されているLUは未使用であると判断され、評価値が0にリセットされる(ステップ3020)。

20

【 0 1 1 8 】

次に、ストレージ管理サーバ10000は、LU評価項目テーブル13400に登録された各項目ごとに、ステップ3020で抽出された全ての未使用LUに評価値を与える。具体例は後述する(ステップ3030)。

【 0 1 1 9 】

その後、ストレージ管理サーバ10000は、LU管理テーブル245'に登録された全てのLUのうち、「評価値」欄2459の値が最小のLUをコピー元LU、又はコピー先LUの候補として出力する(ステップ3040)。

【 0 1 2 0 】

尚、LU管理テーブル245'に登録された評価値は、ストレージ管理サーバ10000が、各ディスクアレイ701に対してLU管理テーブル245の更新を指示するとき(ステップ2240、2440)に、一緒に更新される。

30

【 0 1 2 1 】

以下、ステップ3030の具体例を四つ説明する。

【 0 1 2 2 】

第一の例として、ユーザが「種別」を「FC」と指定したときの評価を考える。ストレージ管理サーバ10000は、「評価対象LU管理テーブル」欄13400bを確認することで、ユーザの指定値「FC」と比較されるべきLU管理テーブル245'の欄は「LU種別」欄であることを確認する。その後、ストレージ管理サーバ10000は、抽出された未使用LUのうち一つを選択して、そのLUに対応するLU管理テーブル245'の「LU種別」欄の内容を確認する。当該「LU種別」欄に登録された情報が「FC」の場合、ユーザが指定した条件と合致するので、ストレージ管理サーバ10000は、「種別」項目に対応する「条件成立時ペナルティ」欄13400dに登録された値に従い、当該未使用LUのLU管理テーブル245'の評価値2459の欄に0を加算する。当該「LU種別」欄に登録された情報が「ATA」の場合、ユーザが入力した条件と合致しないので、ストレージ管理サーバ10000は、「種別」項目に対応する「条件不成立時ペナルティ」欄13400eに登録された値に従い、当該未使用LUのLU管理テーブル245'の評価値2459の欄に+100を加算する。

40

【 0 1 2 3 】

50

ストレージ管理サーバ10000は、上記のような処理、すなわち、各項目ごとにユーザが指定した条件と合致するかどうかを判定し、その結果に応じて所定の値を評価値として各LU管理テーブル245'の評価値欄に加算する処理を、抽出された全ての未使用LUに対してLU評価項目テーブル13400に記載された全ての項目の評価が適用されるまで繰り返す。本処理の結果、全ての未使用LUのLU管理テーブル245'の「評価値」欄2459の値が定まる。この具体例では、ユーザがディスク装置の種別を"FC"と選択することにより、ストレージ管理サーバ10000が、FCディスク装置から構成されるLUを候補として提供することができることを示す。

【0124】

第二の例として、ユーザがバックアップ運用のためのLUを作成するときの評価を考える。データのバックアップを複数世代保持する運用形態では、バックアップ用のLUに、安価なディスク装置から構成されるLUや、余命時間が短いため通常業務運用には利用されにくいディスク装置から構成されるLUを用いることが考えられる。このような場合のために、本例では、図14記載のような、"LU運用"という評価項目を定義する。本評価項目は、ユーザが"LU運用"を"バックアップ"と指定した場合、ストレージ管理サーバ10000が、"LU種別"が"ATA"であり、かつ、"LU余命時間"が"1000時間以下"のLUを優先的に候補として提供するようにLUを評価する項目である。尚、評価手順は第一の例と同一である。この評価項目により、ストレージ管理サーバ10000は、よりバックアップの使用に適したLUの候補をユーザに提示することができる。

【0125】

第三の例として、ユーザの求める信頼性に応じたLU作成をするための評価を考える。このような場合、図14記載のような、"信頼性"という評価項目において"LU余命時間"を評価すると定義する。本評価項目は、ユーザが"信頼性"を"高"と指定した場合、ストレージ管理サーバ10000が、"LU余命時間"が"30000時間以上"のLUを優先的に候補として提供しよう、各LUを評価する項目である。この評価項目により、ストレージ管理サーバ10000は、予め定義された評価基準を用いて余命時間を評価して、適切なLU候補を提供する。

【0126】

第四の例として、ユーザの設定するコピー先の数に応じたLU作成をするための評価を考える。第二の例で述べたデータの複数世代保持のために、リモートコピーや装置内レプリケーションの機能を利用して多数のコピー先を作成する場合がある。このような場合、図14記載のような、"コピー先数"という評価項目において"LU種別"を評価すると定義する。本評価項目は、ユーザが多数のコピー先を指定した場合、ストレージ管理サーバ10000が、安価なATA LUを優先的に探索して候補としてユーザに提示しよう、各LUを評価する項目である。本評価項目により、ストレージ管理サーバ10000は、ユーザが指定したコピー先の数に基づき、予め定義された評価基準を用いてLU種別を評価して、適切なコピー先LU候補を提供する。

【0127】

本実施形態において、ストレージ管理サーバ10000は、入出力I/Fが異なるディスク装置を混在させたディスクアレイ701において、LUレベルでの寿命管理を実現する。以下、この処理をLU寿命管理処理と称する。なお、特に断りのない限り、各処理はストレージ管理サーバ10000のCPU11000がストレージマネージャ13100を実行することで行われる。

【0128】

図20は、LU余命管理処理の手順を示すフローチャートである。

【0129】

ストレージ管理サーバ10000は、ストレージリポジトリ13200で定期的に更新される、ディスクアレイ701A及びBそれぞれのLU管理テーブル245'の情報を用いて、LU余命時間を定期的に監視する。具体的には、ストレージ管理サーバ10000は、LU余命閾値テーブル13500の「装置ID」欄13500bおよび「対象LURIS

10

20

30

40

50

ト」欄 13500c で特定される各 LU について、各 LU に対応する LU 管理テーブル 245' の LU 余命時間 245g に登録された値と、各 LU に設定された閾値（各 LU に対応する「余命閾値」欄 13500d に登録された値）とを比較する。なお、「監視対象 LU リスト」欄の設定状況により、一つの LU に対して複数の閾値が適用できる場合は、「余命閾値」欄の値が最大となる閾値を適用する(ステップ 4010)。

【0130】

比較の結果、「余命閾値」欄 13500d に登録された値に比べて各 LU の余命時間が下回る場合は、ストレージ管理サーバ 10000 は、その LU を警告出力が必要な LU と判断する(ステップ 4020)。

【0131】

ステップ 4020 において警告対象となる LU が存在しなかった場合、ストレージ管理サーバ 10000 は、ステップ 4010 の処理に戻り、LU 余命管理理を続ける。

【0132】

ステップ 4020 において警告対象となる LU が存在した場合、ストレージ管理サーバ 10000 は、該当する警告出力が必要な LU について、余命時間が閾値を下回ったディスク装置を特定し、そのディスク装置について自動移行が指定されているかどうか確認する。具体的には、余命時間が閾値を下回った LU が属する LU 管理テーブル 245' の「ドライブ No. リスト」欄 2458 に記載される全てのディスク装置に対して、ドライブ管理テーブル 240' の「余命時間」欄 2408 を調べ、余命時間が閾値を下回ったディスク装置を特定し、さらに、「自動移行指定」欄 249 を確認する(ステップ 4030)。

【0133】

ステップ 4030 で、余命時間が閾値を下回った LU を構成する全てのディスク装置で自動移行が指定されている場合は、ストレージ管理サーバ 10000 は、警告のみ出力する。ここで、ディスクアレイ装置が自動移行操作を開始していない場合は、自動移行を指示する。警告内容の具体例として、「監視時刻」「ディスクアレイの装置 ID」「LU No.」「余命時間が閾値を下回ったドライブ No.」「関連するコピー元あるいはコピー先 LU の装置 ID 及び LU No.」が挙げられる。

【0134】

なお、「関連するコピー元あるいはコピー先 LU の装置 ID 及び LU No.」の出力に際しては、ストレージリポジトリ 13200 に保持された LU ペア管理テーブル 246 の情報を用いる。「関連するコピー元あるいはコピー先 LU の装置 ID 及び LU No.」を出力することにより、ユーザは、警告対象となった LU が他のコピー元あるいはコピー先 LU に影響を与える可能性を即座に認識し、場合によっては、余命超過に伴う LU ペアの手動変更を実施することができる(ステップ 4040)。

【0135】

ステップ 4030 において、余命時間が閾値を下回った LU を構成するディスク装置のうち、一つでも自動移行が指定されていない場合は、ストレージ管理サーバ 10000 は、ユーザに対して警告を出力すると共に、移行指定あるいはドライブ交換指示を促すメッセージを出力する。警告内容の具体例として、ステップ 4040 記載の内容に加えて、「移行先候補となるドライブ No.」が挙げられる。

【0136】

「移行先候補となるドライブ No.」の決定に際しては、ストレージ管理サーバ 10000 が、ストレージリポジトリ 13200 に保持されたドライブ管理テーブル 240' の情報を用いて、同じドライブ種別でアレイ構成設定を行っていない未使用ディスク装置を探索すること決定する。「移行先候補となるドライブ No.」の出力により、ユーザは、移行指定するための移行先 LU を探す負担を軽減することができる。なお、ディスクアレイがディスク装置の自動移行先を自動的に決定することで、「移行先候補となるドライブ No.」の出力は省略されても良い(ステップ 4050)。

【0137】

本実施形態によれば、入出力 I/F や信頼性が異なるディスクドライブを混在させたディ

10

20

30

40

50

スクアレイのLUを作成するとき、ユーザの指示した特性に基づき、ディスクドライブの性能や信頼性を考慮してLUを作成することができる。また、ディスクアレイがリモートコピーや装置内レプリケーションの機能を具備する場合、そのコピー先LUについても、ユーザの指示した特性に基づき、ディスクドライブの性能や信頼性を考慮して自動的に決定することができる。さらに、ユーザにとって意識しにくいディスクアレイ内のディスク装置の寿命管理を、論理ボリュームであるLUのレベルで行うことができる。

【0138】

なお、本実施形態のストレージ管理サーバ10000は、ディスクアレイ700が第2の実施形態のような構成であっても、ディスク装置の性能や信頼性を考慮してLU作成、およびLUレベルの寿命管理を行うことができる。

10

【0139】

又、本実施形態においてストレージ管理サーバ10000が行っていたLU単位の寿命管理を、第一の実施形態におけるディスクアレイ700の制御部200が実行してもよい。この場合、第一の実施形態で行われていたディスク装置単位の寿命管理におけるパラメータをLU余命に変更すればよい。そして、LU余命がある閾値を下回るLUを発見した場合には、そのLUを構成するディスク装置を検出し、その各々について、自動移行の有無等の確認を行う処理を行えばよい。

【0140】

【発明の効果】

本発明によれば、信頼性、例えば寿命や余命が異なるディスク装置を混在させて、同一のディスクアレイ内で管理することが可能となる。また、アレイを構成するディスク装置の信頼性を考慮して、ボリュームのユーザ用途を決定することができる。また、アレイを構成するディスク装置の性能を考慮して、ボリュームのユーザ用途を決定することができる。また、制御方式の違うディスク装置を混在させて、同一のディスクアレイ内で管理することが可能となる。

20

【0141】

また、ファイバチャネルディスク装置を接続するインターフェースであるファイバチャネルアービトラレーテッドループに、ATAディスク装置を接続することができるという効果がある。

【0142】

さらに、ディスク装置の性能や信頼性を考慮したLUの自動作成、および、LUレベルでの寿命管理を提供し、ユーザの記憶装置システムの管理の負担を軽減することが可能となる。

30

【図面の簡単な説明】

【図1】第1の実施形態におけるシステム構成図である。

【図2】第1の実施形態におけるドライブ管理テーブルを示す図である。

【図3】第1の実施形態における入力部を示す図である。

【図4】第1の実施形態における出力部を示す図である。

【図5】第1の実施形態におけるLU管理テーブルの一例を示す図である。

【図6】第1の実施形態におけるドライブ寿命管理動作のフローチャートである。

40

【図7】第2の実施形態におけるシステム構成図である。

【図8】第2の実施形態におけるポートバイパス回路を示す図である。

【図9】第2の実施形態におけるFC-ATA変換I/Fを示す図である。

【図10】第3の実施形態におけるシステム構成図である。

【図11】第3の実施形態におけるLUペア管理テーブルの一例を示す図である。

【図12】第3の実施形態におけるストレージ管理サーバの構成図である。

【図13】第3の実施形態における装置リストの一例を示す図である。

【図14】第3の実施形態におけるLU評価項目テーブルの一例を示す図である。

【図15】第3の実施形態におけるLU割当処理のフローチャートである。

【図16】第3の実施形態におけるホストLU割当処理のフローチャートである。

50

【図17】第3の実施形態におけるコピー先LU割当処理のフローチャートである。

【図18】第3の実施形態における割当候補LU探索処理のフローチャートである。

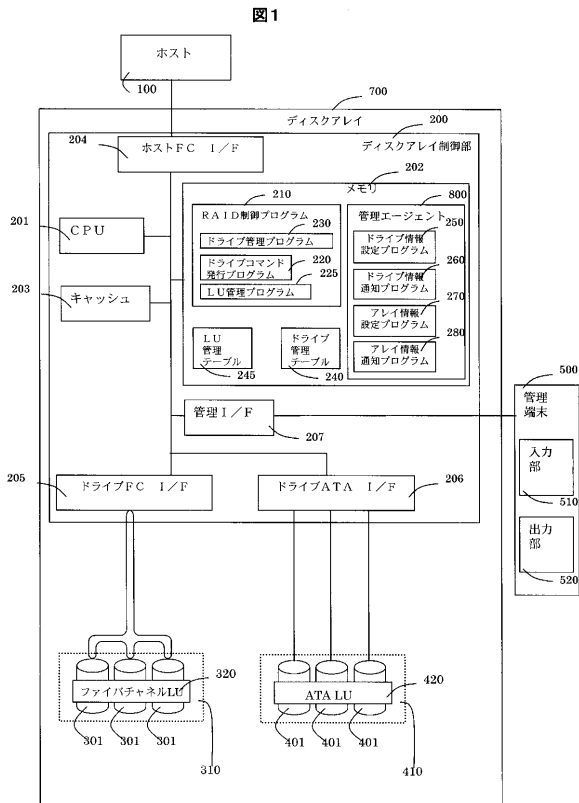
【図19】第3の実施形態におけるLU余命閾値テーブルの一例を示す図である。

【図20】第3の実施形態におけるLU寿命管理処理のフローチャートである。

【符号の説明】

100...ホスト、200...ディスクアレイ制御部、310...ファイバチャネルディスク群、410...ATAディスク群、500...管理端末、700...ディスクアレイ、10000...ストレージ管理サーバ

【図1】

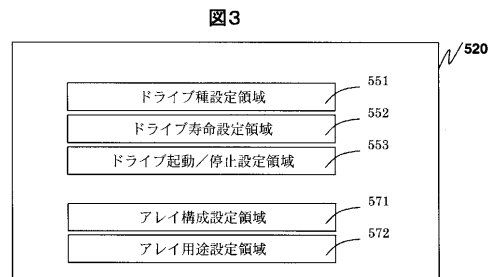


【図2】

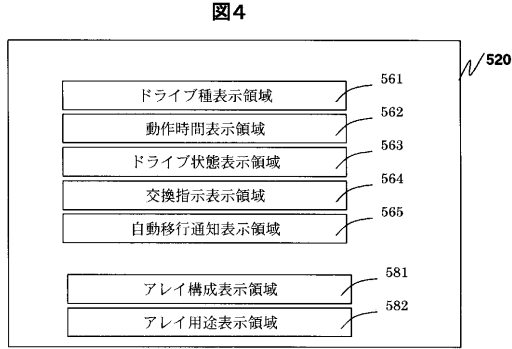
図2

ドライブNo.	ドライブ種別	アレイ構成	用途	起動状況	累積時間	寿命設定	余命時間	自動移行指定
0	FC	(1)RAID5	DB	ON	10000hr	40000hr	30000hr	0
1	FC	(1)RAID5	DB	ON	10000hr	40000hr	30000hr	0
2	FC	(1)RAID5	DB	ON	10000hr	40000hr	30000hr	0
3	ATA	(2)RAID5	FS	ON	23000hr	25000hr	2000hr	1
4	ATA	(2)RAID5	FS	ON	23000hr	25000hr	2000hr	1
5	ATA	(2)RAID5	FS	ON	23000hr	25000hr	2000hr	1
2401	N	N	N	N	N	N	N	N
2402	N	N	N	N	N	N	N	N
2403	N	N	N	N	N	N	N	N
2404	N	N	N	N	N	N	N	N
2405	N	N	N	N	N	N	N	N
2406	N	N	N	N	N	N	N	N
2407	N	N	N	N	N	N	N	N
2408	N	N	N	N	N	N	N	N
2409	N	N	N	N	N	N	N	N

【図3】



【図4】



【図5】

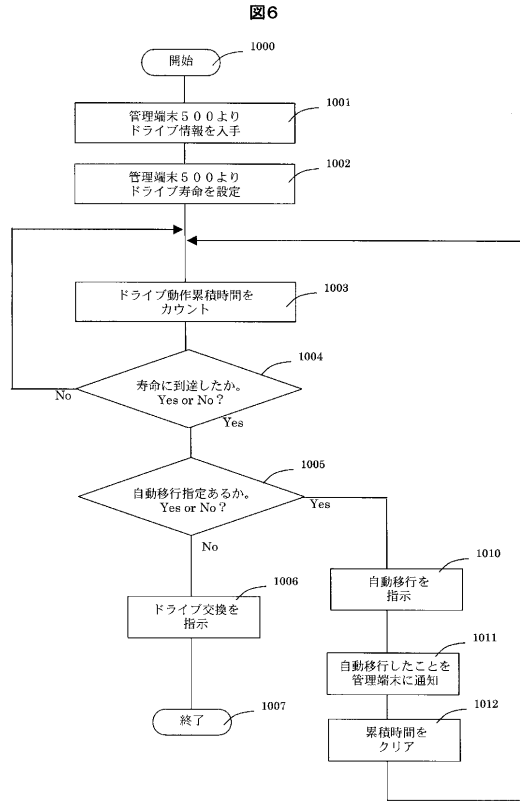
図5

245

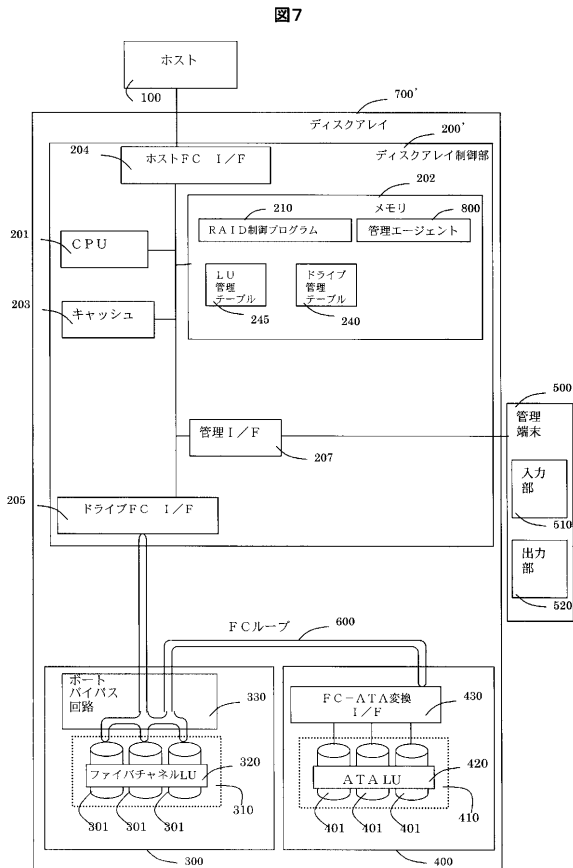
LU No.	ホスト割当状況	SCSI ID	LUN	容量	LU 種別	LU余命時間	ドライブNo. リスト	評価値
0	有	2	1	50GB	FC	40000hr	0, 1, 2	9999
1	無	-	-	50GB	FC	40000hr	0, 1, 2	0
2	有	3	1	50GB	ATA	10000hr	3, 4, 5	9999
3	無	-	-	50GB	ATA	10000hr	3, 4, 5	0

2451 2452 2453 2454 2455 2456 2457 2458 2459

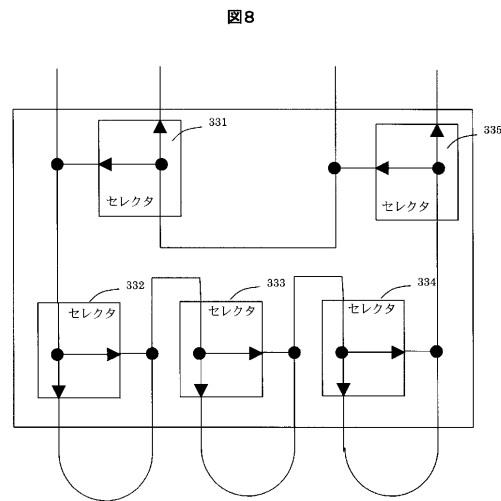
【図6】



【図7】

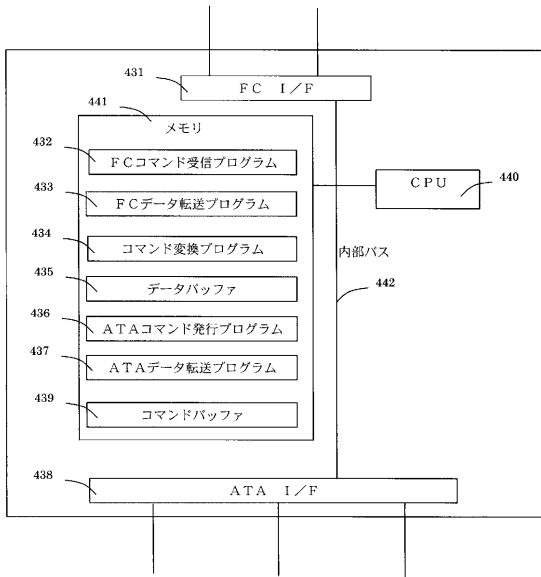


【図8】



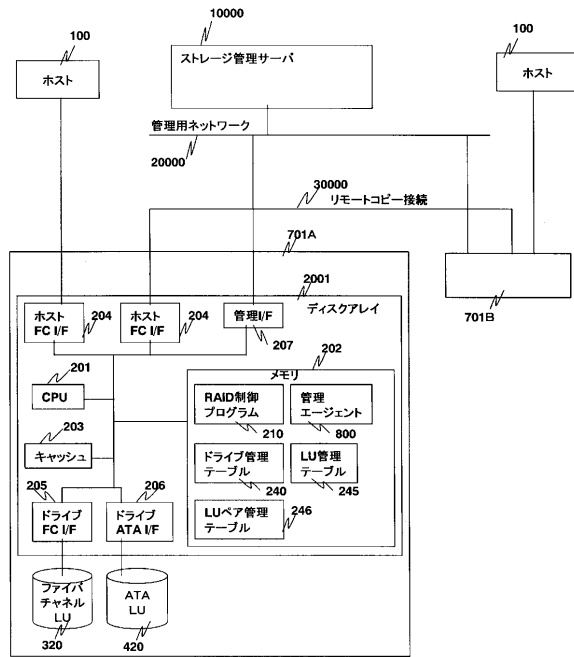
【図9】

図9



【図10】

図10



【図11】

図11

LUベア No.	コピー元LU情報			コピー先LU情報			ベア 状態
	LU No.	ベア 属性	LU 種別	装置 ID	LU No.	LU 種別	
0	100	ローカル	FC	ディスクアレイA	200	FC	Pair
1	110	ローカル	FC	ディスクアレイA	210	ATA	Pair
2	120	ローカル	ATA	ディスクアレイA	220	ATA	Split
3	130	リモート	FC	ディスクアレイB	230	FC	Pair
4	140	リモート	ATA	ディスクアレイB	240	ATA	Split

246a 246b 246c 246d 246e 246f 246g 246h

【図14】

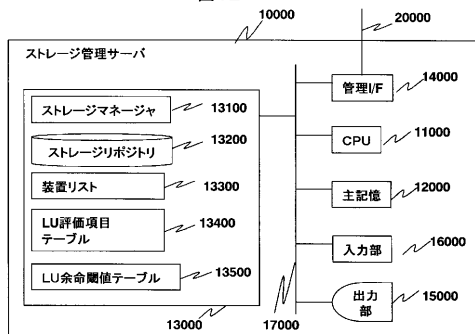
図14

評価項目	評価対象 LU管理テーブル欄	入力条件 読み替え	条件成立時 ペナルティ	条件不成立時 ペナルティ
種別	LU種別	無し	0	+100
容量	容量	無し	GBあたり+2	+99999
用途	LU種別	"DB"=FC, "FS"=ATA	0	+10
余命	LU余命時間	無し	0	1hrあたり+100
性能	LU種別	90MB/s以上でFC	0	+50
価格	LU種別	\$/GB以下でATA	0	+30
LU運用	LU種別	"バックアップ"=ATA	0	+1000
LU運用	LU余命時間	"バックアップ"=1000hr以下	0	1hrあたり+1000
信頼性	LU余命時間	"高"=30000hr以上 "中"=15000hr以上	0	+1000
3C先数	LU種別	"1~5"=コピー元と同じ "6以上"=ATA	0	+1000

13400a 13400b 13400c 13400d 13400e

【図12】

図12



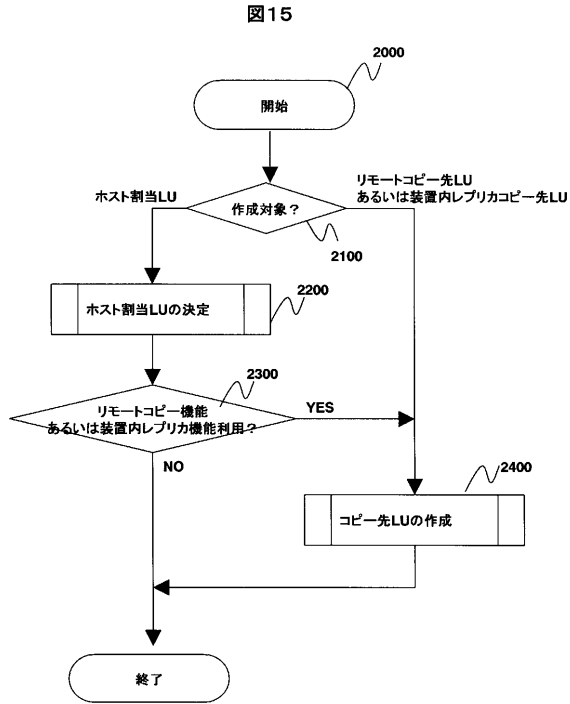
【図13】

図13

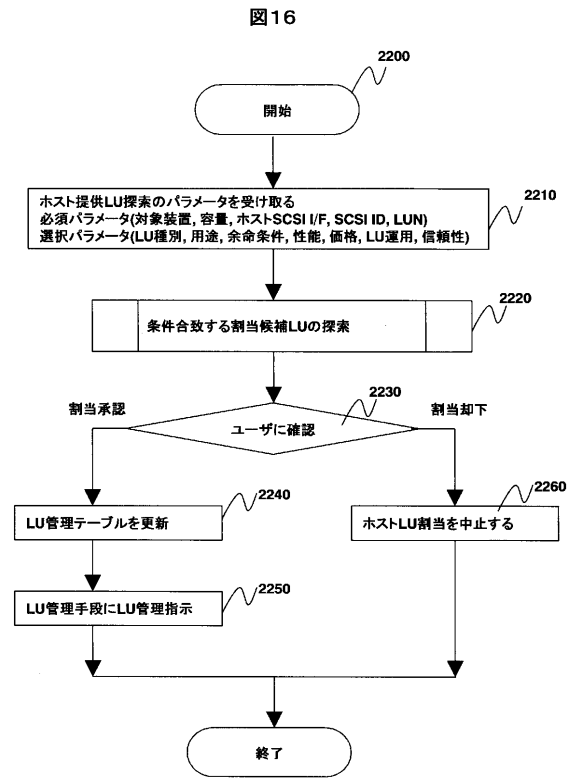
対象装置No.	装置ID	IPアドレス
1	ディスクアレイA	100.100.100.100
2	ディスクアレイB	100.100.100.101

13300a 13300b 13300c

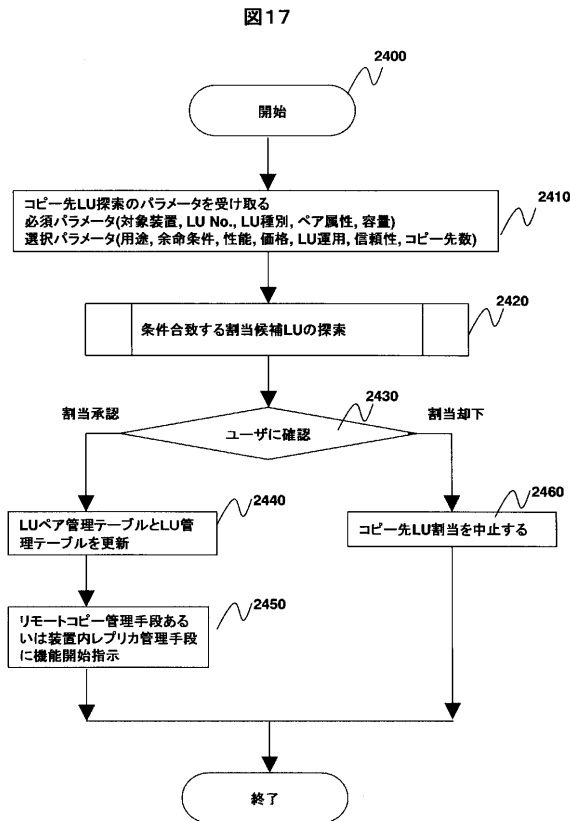
【 図 1 5 】



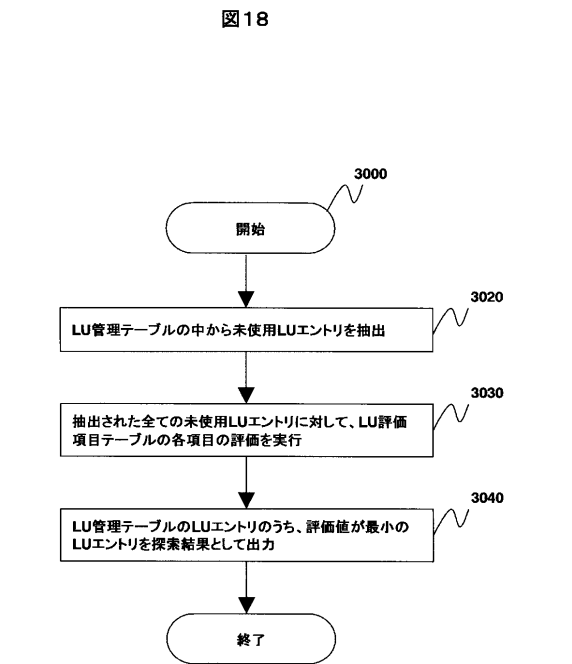
【 図 1 6 】



【 図 1 7 】



【 図 1 8 】



【図19】

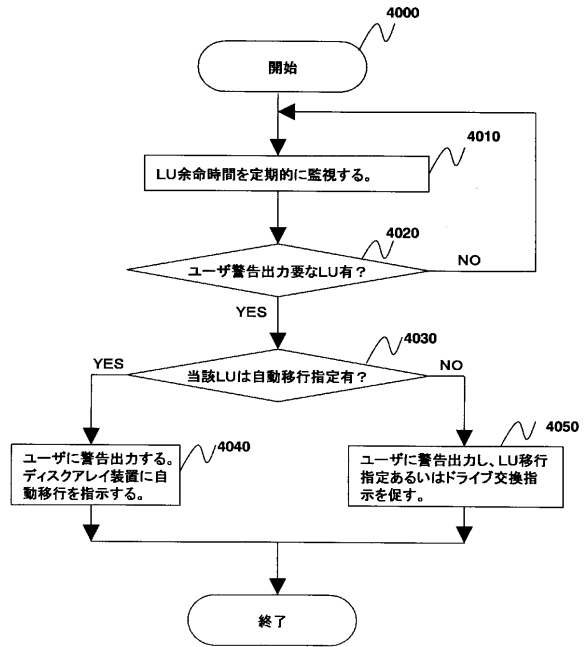
図19

図値No.	装置ID	監視対象LUリスト	余命閾値
1	ディスクアレイA	全LU	100hr
2	ディスクアレイB	全LU	500hr
3	ディスクアレイB	0, 1, 2, 3, 4, 5	1000hr

13500a 13500b 13500c 13500d

【図20】

図20



フロントページの続き

- (72)発明者 萬年 暁弘
神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内
- (72)発明者 山本 政行
神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

合議体

- 審判長 大野 克人
審判官 清水 稔
審判官 安島 智也

- (56)参考文献 特開2000 - 200156 (JP, A)
特開2001 - 67187 (JP, A)
特開2000 - 148705 (JP, A)
特開平3 - 103956 (JP, A)
特開2000 - 339098 (JP, A)
特開2002 - 222061 (JP, A)
“Netfinity Server Disk Subsystems 日本語版”
，日本アイ・ビー・エム株式会社，2000年 6月30日，第4版，p. 79 ~ 90

- (58)調査した分野(Int.Cl. , DB名)
G06F 3/06, G11B 27/00