

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2005-293205

(P2005-293205A)

(43) 公開日 平成17年10月20日(2005.10.20)

(51) Int. Cl.⁷

G06F 3/06

F I

G06F 3/06 3 O 2 A
 G06F 3/06 3 O 1 S
 G06F 3/06 5 4 O

テーマコード (参考)

5 B O 6 5

審査請求 未請求 請求項の数 10 O L (全 19 頁)

(21) 出願番号 特願2004-107030 (P2004-107030)
 (22) 出願日 平成16年3月31日 (2004.3.31)

(71) 出願人 000005223
 富士通株式会社
 神奈川県川崎市中原区上小田中4丁目1番
 1号
 (74) 代理人 100074099
 弁理士 大菅 義之
 (74) 代理人 100067987
 弁理士 久木元 彰
 (72) 発明者 大村 英明
 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
 Fターム(参考) 5B065 BA01 CA30 CE11 CH01 EA23

(54) 【発明の名称】 記憶制御装置、制御方法、および制御プログラム。

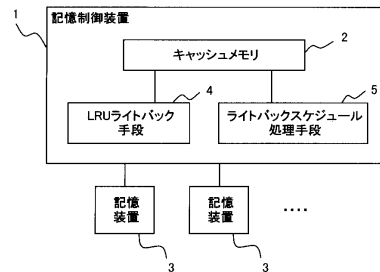
(57) 【要約】

【課題】 レベルの異なる複数のRAID装置へのデータのライトバックを、一部のRAID装置に偏ることなく実行する。

【解決手段】 RAID制御装置1内のキャッシュメモリ2に格納されているデータの複数のRAID装置3へのライトバックをLRU法によって行う手段4と、手段4によって実行されているライトバックの数が少ない装置3を選択して、選択した装置3に対してデータのライトバックを行う手段5とを備える。

【選択図】 図1

本発明の記憶制御装置の原理構成ブロック図



【特許請求の範囲】**【請求項 1】**

複数の記憶装置を制御する記憶制御装置において、
該記憶制御装置内のキャッシュメモリに格納されているデータの前記複数の記憶装置へのライトバックをLRU法によって行うLRUライトバック手段と、
該LRUライトバック手段によって実行されているライトバックの数が少ない記憶装置を選択して、該選択した記憶装置に対するデータのライトバックを行うライトバックスケジュール処理手段とを備えることを特徴とする記憶制御装置。

【請求項 2】

前記記憶制御装置において、
前記複数の記憶装置を、該記憶装置が実行しているライトバックの数によって複数のグループに分割した結果を記憶するライトバックグループ管理手段をさらに備え、
前記ライトバックスケジュール処理手段が、ライトバックグループ管理手段の記憶内容に従ってライトバックの数が少ないグループ内の記憶装置を選択することを特徴とする請求項 1 記載の記憶制御装置。

10

【請求項 3】

前記記憶制御装置において、
前記複数の記憶装置毎に実行されているライトバックの数を記憶する記憶装置管理手段をさらに備え、
前記ライトバックスケジュール処理手段が、該記憶装置管理手段の記憶内容に従ってライトバックの数が少ない記憶装置を選択することを特徴とする請求項 1 記載の記憶制御装置。

20

【請求項 4】

前記記憶制御装置において、
該記憶制御装置内のキャッシュメモリの容量を分割して、前記複数の記憶装置に割当てるキャッシュメモリ割当手段と、
該割当てられた容量の中で該キャッシュメモリにデータが格納されていない未使用領域の容量を記憶装置毎に記憶する記憶装置管理手段とをさらに備え、
前記ライトバックスケジュール処理手段が、該記憶装置管理手段の記憶内容に従って該未使用領域の少ない記憶装置を選択することを特徴とする請求項 1 記載の記憶制御装置。

30

【請求項 5】

前記記憶制御装置において、
前記記憶装置に割当てられたキャッシュメモリの容量の内部に格納されているデータのキューを記憶装置毎に管理するキャッシュメモリ管理手段をさらに備え、
前記ライトバックスケジュール処理手段が、選択した記憶装置のキューにおけるLRU位置のデータのライトバックを行うことを特徴とする請求項 4 記載の記憶制御装置。

【請求項 6】

自装置内のキャッシュメモリに格納されているデータの複数の外部の記憶装置へのライトバックをLRU法によって行うLRUライトバック手段と、
該LRUライトバック手段によって実行されているライトバックの数が少ない記憶装置を選択し、該選択した記憶装置に対するデータのライトバックを行うライトバックスケジュール処理手段とを備える記憶制御装置と、
該記憶制御装置によって制御される複数の記憶装置とによって構成されることを特徴とするデータ記憶システム。

40

【請求項 7】

前記記憶装置が複数のディスクを備えるRAID装置であり、
前記記憶制御装置が該複数のRAID装置を制御するRAID制御装置であることを特徴とする請求項 6 記載のデータ記憶システム。

【請求項 8】

複数の記憶装置を制御する記憶制御方法において、

50

キャッシュメモリに格納されているデータの前記複数の記憶装置側へのライトバックをLRU法によって実行し、

該LRU法によって実行されているライトバックの数が少ない記憶装置を選択して、該選択した記憶装置に対するデータのライトバックを実行することを特徴とする記憶制御方法。

【請求項9】

複数の記憶装置を制御する計算機によって使用されるプログラムであって、

キャッシュメモリに格納されているデータの前記複数の記憶装置側へのライトバックをLRU法によって実行する手順と、

該LRU法によって実行されているライトバックの数が少ない記憶装置を選択して、該選択した記憶装置に対するデータのライトバックを実行する手順とを計算機に実行させるためのプログラム。

10

【請求項10】

複数の記憶装置を制御する計算機によって使用される記憶媒体であって、

キャッシュメモリに格納されているデータの前記複数の記憶装置側へのライトバックをLRU法によって実行するステップと、

該LRU法によって実行されているライトバックの数が少ない記憶装置を選択して、該選択した記憶装置に対するデータのライトバックを実行するステップとを計算機に実行させるプログラムを格納した計算機読出し可能可搬型記憶媒体。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明はデータ記憶方式に係り、さらに詳しくはディスクアレイを用いた磁気ディスクシステムに対応して、ホストコンピュータからの指示によってデータのリード/ライトが行われるキャッシュメモリを備えた磁気ディスク制御装置に関する。

【背景技術】

【0002】

近年、データ記録の不揮発性、大容量性などの特徴を持つ磁気ディスク装置、光ディスク装置などのディスク装置がコンピュータの外部記憶装置として広く用いられている。このような磁気ディスク装置等においてデータの内容を二重化したり、パリティデータなどの冗長な情報を付け加えて記録したりするためのディスク装置としてRAID（リダンダント・アレイ・オブ・インエクスペンシブ・ディスク）装置が用いられている。

30

【0003】

このようなRAID装置と、例えばホストコンピュータとの間でのデータの転送を制御するRAID制御装置は大容量のキャッシュメモリを搭載し、そのキャッシュメモリに一時データを保持することによってデータ転送の性能向上を図っている。またこのような制御装置の配下には、扱うべきデータ量の増大に伴って膨大な数の多種多量のRAID装置、すなわちRAID0からRAID5までのレベルのRAID装置が多数接続されて使用される。

【0004】

図14はこのようなRAID制御装置におけるキャッシュデータ管理方式の従来例の説明図である。同図においてRAID制御装置のキャッシュメモリ内のデータは、ホストコンピュータによる使用頻度の最も高いMRUの位置のデータから、最も使用頻度の低いLRUの位置のデータまで、例えばリンクを用いて並べられ、ホストコンピュータからのアクセスがあった場合にはキャッシュヒットしたものとして、最も使用頻度の高いデータとしてMRUの位置に接続される。

40

【0005】

図14においてホストコンピュータからキャッシュメモリに書き込まれたデータに対してはライトバック、すなわち非同期でのディスク装置への書き出しが行われる。このRAID装置（ディスク）へのデータのライトバックは、データ記憶システムとしての全体的

50

な性能を考慮して、LRU法によってアクセス頻度の低いデータ、すなわちLRUの位置のデータから実行される。

【0006】

しかしキャッシュメモリからRAID装置(ディスク)への書き出し性能は、RAIDタイプ(レベル)によって異なる。すなわちデータを複数のディスク装置に分散して書き込むRAID0+1、RAID5では複数のディスクが並列に動作するためにスループット性能が高く、逆に2台のディスク装置にそれぞれ同じ内容のデータを書き込むRAID1ではディスクの性能以上の能力が出せないため、スループット性能が低くなることになる。

【0007】

このようにライトバック性能が異なる複数のRAID装置に書き込むべきデータを1つのキャッシュメモリを介するデータ転送という形式で管理する場合、ライトバック性能が悪いRAID1の装置へのライトデータがキャッシュメモリ上で滞留しやすくなる。キャッシュメモリがそのような滞留データで埋め尽くされてしまうと、それ以降はRAID1のディスク装置へのライトバック性能に引きずられ、スループット性能が良いRAID0+1、RAID5に対するアクセスにおいても性能低下が起こるといった問題点があった。

【0008】

このようなRAIDデータ記憶システムに関する従来技術として次のような文献がある。

【特許文献1】特開平11-288387号公報 「ディスクキャッシュ装置」

【特許文献2】特開2003-196032号公報 「ストレージ装置のライトキャッシュ制御方法及びストレージ装置」

【0009】

特許文献1では、ディスクキャッシュ機能の最適化を図るために、例えばRAID5の動作でもディスクリードを必要とせずにパリティを計算できるRAID3的なアクセスを可能とするように、キャッシュブロック上に有効データを準備しておくディスクキャッシュ装置が開示されている。

【0010】

特許文献2には、ライトバックキャッシュ方式のストレージ装置において、ホスト装置から指定される書込み対象のアドレスが連続していない場合でも、シーケンシャルなデータの書き出しをもれなく実行することができるライトキャッシュ制御方法を用いてスループットを改善する技術が開示されている。

【0011】

しかしながらこのような従来技術においても、RAID装置として各種のレベルの装置が混在する場合に、ライトバック性能が悪いRAID装置の性能に引きずられてデータ記憶システムとしての全体的性能が低下するという問題点を解決することができない。

【発明の開示】

【発明が解決しようとする課題】

【0012】

本発明の課題は、上述の問題点に鑑み、レベルが異なる複数のRAID装置がRAID制御装置の配下に存在する場合にも、できるだけ多くのRAID装置へのライトバックをスケジュール可能とすることによって、データ記憶システムとしての全体的な性能を向上させることである。

【課題を解決するための手段】

【0013】

図1は本発明の記憶制御装置、例えばRAID制御装置の原理構成ブロック図である。同図において記憶制御装置1はキャッシュメモリ2を備え、複数の記憶装置3、例えばRAID装置を制御するものであり、基本的構成要素としてLRUライトバック手段4とライトバックスケジュール処理手段5とを少なくとも備える。

【0014】

10

20

30

40

50

L R Uライトバック手段4は、記憶制御装置1の内部のキャッシュメモリ2に格納されているデータの複数の記憶装置3側へのライトバックを、L R U法(アルゴリズム)によって行うものであり、ライトバックスケジュール処理手段5はL R Uライトバック手段4によって実行されているライトバックの数が少ない記憶装置3を選択して、選択した記憶装置3に対するデータのライトバックを行うものである。

【0015】

発明の実施の形態においては記憶制御装置1が、複数の記憶装置3をその記憶装置3が実行しているライトバックの数によって複数のグループに分割した結果を記憶するライトバックグループ管理手段、例えばライトバックグループ管理テーブルをさらに備え、ライトバックスケジュール処理手段5がライトバックグループ管理手段の記憶内容に従ってライトバックの数が少ないグループ内の記憶装置3を選択することもできる。

10

【0016】

また実施の形態においては、複数の記憶装置3毎に実行されているライトバックの数を記憶する記憶装置管理手段、例えばR A I D装置毎の管理テーブルをさらに備え、ライトバックスケジュール処理手段5が記憶装置管理手段の記憶内容にしたがってライトバックの数が少ない記憶装置3を選択することもできる。

【0017】

さらに実施の形態においては、記憶制御装置内のキャッシュメモリの容量を分割して、複数の記憶装置3に割当てられたキャッシュメモリ割当手段と、割当てられた容量の中でキャッシュメモリにデータが格納されていない未使用領域の容量を記憶装置毎に記憶する記憶装置管理手段、例えばR A I D装置毎の管理テーブルとをさらに備え、ライトバックスケジュール処理手段5が記憶装置管理手段の記憶内容にしたがって未使用領域の少ない記憶装置3を選択することもできる。

20

【0018】

この場合、記憶装置3に割当てられたキャッシュメモリの容量の内部に格納されているデータのキューを記憶装置3毎に管理するキャッシュメモリ管理手段、例えばキャッシュメモリ管理テーブルをさらに備え、ライトバックスケジュール処理手段5が選択した記憶装置3に対応するキューにおけるL R Uの位置のデータのライトバックを行うこともできる。

【0019】

また発明の実施の形態においては、前述の記憶装置3が複数のディスクを備えるR A I D装置であり、記憶制御装置1が複数のR A I D装置を制御するR A I D制御装置であることもできる。

30

【0020】

次に本発明のデータ記憶システムは、自装置内のキャッシュメモリに格納されているデータの複数の記憶装置側へのライトバックL R U法によって行うL R Uライトバック手段と、L R Uライトバック手段によって実行されているライトバックの数が少ない記憶装置を選択し、選択した記憶装置に対するデータのライトバックを行うライトバックスケジュール処理手段とを備える記憶制御装置と、その記憶制御装置によって制御される複数の記憶装置とによって構成される。

40

【0021】

本発明の記憶制御方法においては、記憶制御装置内のキャッシュメモリに格納されているデータの複数の記憶装置側へのライトバックをL R U法によって実行し、L R U法によって実行されているライトバックの数が少ない記憶装置を選択して、選択した記憶装置に対するデータのライトバックを実行する方法が用いられる。

【0022】

またこの方法を計算機によって実行するためのプログラム、およびそのプログラムを格納した計算機読み出し可能可搬型記憶媒体が用いられる。

以上のように本発明によれば、まず記憶制御装置、たとえばR A I D制御装置内のキャッシュメモリを複数の記憶装置、例えばR A I D装置に対応して分割して割当てること

50

よって、キャッシュメモリの内部がスループット性能の悪い記憶装置に掃き出すべきデータで埋め尽くされることが防止される。

【0023】

またこのRAID装置毎に割当てられたキャッシュメモリ容量の中で未使用のメモリ容量を管理し、未使用領域が少ないRAID装置に対するライトバックを実行することによって、キャッシュメモリの空き容量を確保することができる。

【0024】

さらにRAID装置毎にライトバックの実行数を管理するとともに、ライトバックの実行数によってRAID装置をグループ化することによって、ライトバック実行数の少ないRAID装置の検索が容易となり、多くのRAID装置に対してライトバックをスケジューリングすることが可能となる。

10

【発明の効果】

【0025】

本発明によれば、RAID制御装置の配下にRAIDタイプ(レベル)の異なる複数のRAID装置が存在する場合にも、ライトバック性能が悪いRAID装置による性能低下を防止し、データ記憶システムとしての全体性能を向上させることが可能となり、RAID記憶装置システムにおける性能向上に寄与するところが大きい。

【発明を実施するための最良の形態】

【0026】

図2は本発明のRAID制御装置が用いられるデータ記憶システム全体構成ブロック図である。同図においてシステムはRAID制御装置10、ホストコンピュータ(サーバ)11、および複数のRAID装置12によって構成されている。

20

【0027】

RAID制御装置10は、ホストコンピュータ11側との間のデータ転送を制御するためのチャンネルアダプタ13、RAID装置12との間でデータ転送を制御するデバイスアダプタ14、およびRAID制御装置10の全体を制御する中央制御部15を備えている。

【0028】

中央制御部15は、ホストコンピュータ11とRAID装置12との間で転送が行われるデータを一時的に保持するキャッシュメモリ16、各種の制御情報を格納するテーブルとしてのキャッシュメモリ管理テーブル17、RAID装置毎の管理テーブル18、およびライトバックグループ管理テーブル19を備えている。ここでキャッシュメモリ16は、例えば2GBの搭載容量を持つものとする。

30

【0029】

図2においてRAID装置12は、RLU(レイド・ロジカル・ユニット)#00からRLU#05までの6台があるものとし、このうちRLU#04、およびRLU#05のレベルはRAID1である。RLU#00とRLU#01とのレベルはRAID5であり、RLU#00はデータ格納用のディスク4台とパリティ格納用の1台のディスクを備え、RLU#01はデータ格納用の15台、パリティ格納用の1台のディスクを備えるものとする。RLU#02、およびRLU#03のレベルはRAID0+1であり、RLU#02はデータ格納用の4台、ミラー用の4台のディスクを備え、RLU#03はデータ格納用の8台とミラー用の8台のディスクを備えるものとする。

40

【0030】

本実施形態においては、まずRAID制御装置10の内部のキャッシュメモリ16をRAID装置12のスループット性能に対応して分割し、分割結果を各RAID装置12に割り付けることによって、キャッシュメモリ16がスループット性能の悪いRAID装置にライトバックすべきデータで埋め尽くされることが防止するものとする。図3はこのキャッシュメモリ分割割付処理のフローチャートである。同図において装置の電源がオンされ、装置立上げ処理の開始に伴ってキャッシュメモリ分割処理が初期化処理の1つとして実行される。

50

【0031】

図3のステップS1においてRAID制御装置10のキャッシュメモリ容量、ここでは2GBが獲得され、それがキャッシュメモリ管理テーブル17に設定され、ステップS2でRAID制御装置10内の構成情報が読み込まれ、ステップS3でその構成情報からRAID制御装置10の配下の複数のRAID装置12に関する情報が獲得される。そしてその情報に対応して複数のRAID装置12に対してそれぞれステップS4からステップS9の処理が実行される。

【0032】

まずステップS4で処理を行うべきRAID装置12に関する情報が存在するか否かが判定され、存在する場合にはステップS5でRAIDタイプ(レベル)とディスクの数が獲得され、ステップS6でRAIDのタイプがRAID5であるか否かが判定され、5である場合にはステップS7で割当てべきキャッシュ容量が64MB×ディスク数として計算され、RAID5でない場合には、ステップS8で128MB×ディスク数としてキャッシュ容量が計算された後に、ステップS9でそのRAID装置毎の管理テーブル18に求められたキャッシュ容量が設定されて、ステップS4以降の処理が繰り返される。全てのRAID装置12に対する処理が終了した時点でキャッシュメモリ分割処理を終了する。

10

【0033】

図4はキャッシュメモリ使用時処理のフローチャートである。同図は、例えば図2のRAID制御装置10の内部のキャッシュメモリ16に、ホストコンピュータ11側からデータが格納される場合の制御情報(テーブル)を利用した処理のフローチャートである。

20

【0034】

図4において処理が開始されると、まずステップS11で図2のキャッシュメモリ管理テーブル17からキャッシュメモリ16内の未使用のキャッシュメモリ容量が獲得され、ステップS12で未使用のキャッシュメモリ、すなわちデータを格納できるキャッシュメモリの容量があるか否かが判定され、ない場合には直ちに処理を終了する。

【0035】

未使用のキャッシュメモリがある場合には、ステップS13でRAID装置毎の管理テーブル18から未使用のキャッシュメモリ容量が獲得される。この処理では、図3においてそれぞれのRAID装置12に対して設定されたキャッシュ容量、例えばアドレスによって指定されるキャッシュ容量に対応して、データを格納すべきアドレスに対応するRAID装置毎の管理テーブル18から未使用のキャッシュメモリ容量が獲得されるが、次のステップS14でそのアドレスに対応するRAID装置毎の管理テーブル18において未使用のキャッシュメモリがあるか否かが判定され、未使用のキャッシュメモリがない、すなわちそのアドレスにデータが格納できない場合にも直ちに処理を終了する。

30

【0036】

ステップS14で未使用のキャッシュメモリがある場合には、ステップS15でその未使用のキャッシュメモリに、例えばデータが格納されるとともにそのキャッシュメモリを使用中状態とする処理が行われ、ステップS16でキャッシュメモリ管理テーブル17の未使用キャッシュメモリ容量から使用中状態とされた容量が減算され、ステップS17でRAID装置毎の管理テーブル18に格納されている未使用キャッシュメモリ容量、すなわちそのRAID装置12に割当てられたキャッシュメモリ容量の中で使用されていないキャッシュメモリ容量から使用中状態とされた容量が減算されて処理を終了する。なおRAID装置毎の管理テーブルの具体的な内容についてはさらに図6で後述する。

40

【0037】

図5はキャッシュメモリ解放時のフローチャートである。この処理は、例えばキャッシュメモリからデータが削除される場合に実行される処理であり、まずステップS21で解放されるキャッシュメモリのアドレスが獲得され、ステップS22でこのキャッシュメモリが未使用状態とされ、例えばキャッシュメモリ管理テーブル17に格納されている未使用キュー(リンク)にそのアドレスのキャッシュメモリが接続されて解放が行われ、ステ

50

ップS 2 3でキャッシュメモリ管理テーブル1 7の未使用キャッシュメモリ容量に解放された容量が加算され、ステップS 2 4でRAID装置毎の管理テーブル1 8内の同じく未使用キャッシュメモリ容量に解放された容量が加算されて処理を終了する。

【0038】

図6は図2のRAID装置毎の管理テーブル1 8とライトバックグループ管理テーブル1 9の格納内容の具体例の説明図である。RAID装置毎の管理テーブル1 8は、RLU(RAID ロジカル・ユニット)管理テーブルとも呼ばれ、各RAID装置1 2、ここではRLU# 0 0からRLU# 4 0 9 5までの全部で4 0 9 6台のRAID装置1 2に割当てられたキャッシュメモリのページ数、割当てられたページの中で未使用のページ数としてのページレフト、およびライトバックの実行数がそれぞれのRAID装置1 2に対して格納されている。

10

【0039】

ライトバックグループ管理テーブル1 9は、ライトバックの実行数が同じくらいのRAID装置をグループ化して管理するものであり、ここではグループAはキャッシュメモリが未使用でライトバックが不必要なRAID装置群、グループBはキャッシュメモリが使用され、ライトバックの実行数が0のRAID装置群、グループCはキャッシュメモリが使用され、ライトバックの実行数が1から5までのRAID装置群、グループDはキャッシュメモリが使用され、ライトバックの実行数が6以上のRAID装置群であり、図6では各グループからリンクによって接続される形式でRAID装置毎の管理テーブル1 8が管理される。ここではグループB、すなわちRLU# 0 2とRLU# 4 0 9 5においてライトバックの実行数が少なく、また未使用ページ容量が少ないことからライトバックが必要であることが容易に判定される。

20

【0040】

図7は本実施形態におけるライトバック実行処理のフローチャートである。本実施形態では、前述のようにまず通常のLRU法によるライトバックスケジュールによってライトバックの実行処理が行われた後に、ライトバックの実行数の少ないRAID装置がチェックされ、RAID装置単位にさらにライトバックスケジュールが実行されるが、図7のライトバック実行処理は、通常のLRU法によるライトバックにおいても、またRAID装置単位のライトバックにおいても同様に用いられるものである。

【0041】

図7において処理が開始されると、まずステップS 3 1でライトバック対象のキャッシュデータが獲得され、ステップS 3 2で図2のデバイスアダプタ1 4に対してライトバックデータのキャッシュアドレスが通知され、ステップS 3 3でそのデータのライトバックが指示される。そしてステップS 3 4でライトバックすべきキャッシュデータに対応するRLUの番号からRAID装置毎の管理テーブル1 8が求められ、ステップS 3 5でそのテーブルのライトバック実行数に1が加算され、ステップS 3 6で加算された実行数に対応してライトバックグループ管理テーブル1 9におけるグループが求められる。

30

【0042】

続いてステップS 3 7でライトバック実行数に1を加算した結果、RAID装置1 2が現在のグループと同一グループに留まっているか否かが判定され、同一グループである場合には直ちに処理を終了し、グループの変更を行うべき場合には、ステップS 3 8で現在のグループからそのRAID装置毎の管理テーブル1 8が切断され、ステップS 3 9でその管理テーブル1 8が新しいグループに接続されて処理を終了する。

40

【0043】

図8は、図2のデバイスアダプタ1 4からライトバックの実行処理が完了したことを示す通知を受信した後に行われるライトバック完了処理のフローチャートである。実際にはキャッシュメモリ1 6からRAID装置1 2へのデータのライトバックにはかなりの時間を要するため、図7においてデバイスアダプタ1 4にデータのライトバックが指示されてから、図8のライトバック完了処理が開始されるまでには、かなりの時間が必要となる。

【0044】

50

図8のステップS41でライトバック対象のキャッシュデータが獲得され、ステップS42でそのライトバックデータがRAID装置毎のキャッシュデータのキューから切断される。このRAID装置毎のキャッシュデータのキューは図9に示される。

【0045】

図9においてMRUの位置、すなわち最も使用頻度の高いデータからLRUの位置、すなわち最も使用頻度の低いデータまでのそれぞれのデータに対してRAID装置毎のキュー（リンク）が張られている。同図において実線の矢印は、RAID1のRAID装置のデータのリンク、一点鎖線は、RAID5の装置のデータのリンク、点線はRAID0+1の装置のデータのリンク（それぞれ双方向）を示し、上向きの矢印は、キューの先頭側のキャッシュデータを示し、その矢印をプリビアスポインタと呼び、下の矢印は末尾側のキャッシュデータを示し、その矢印をネクストポインタと呼ぶことにする。なお、図2では各タイプのRAID装置が2台ずつ存在するため、それぞれの装置に対応して、例えば実線のキュー（リンク）は2本張られることになる。

10

【0046】

図8のステップS42でライトバックデータがキューから切断された後に、ステップS43でライトバックが行われたキャッシュデータに対応するRLU番号からRAID装置毎の管理テーブル18が求められ、ステップS44でその管理テーブル18内のライトバック実行数がデクリメントされ、ステップS44でデクリメントされた結果のライトバック実行数からライトバックグループ管理テーブル19におけるグループが図7のステップS36と同様に求められ、その後ステップS46からS48で図7のステップS37からS39と同様の処理が実行されて処理を終了する。

20

【0047】

図10、および図11は、このRAID装置毎のキューの管理処理のフローチャートであり、図10はキュー接続処理のフローチャート、図11はキュー切断処理のフローチャートである。図10のキュー接続処理が開始されると、まずステップS51でキューのMRUポインタが獲得され、ステップS52でこのポインタが存在したか否かが判定され、存在している場合にはキャッシュデータがあることになり、ステップS53でMRUポインタの示すキャッシュデータの前に今回新たなキャッシュデータを接続するためにプリビアスポインタを新たなキャッシュデータのアドレスにセットし、ステップS54で現在のMRUポインタの示すキャッシュデータを今回接続すべきキャッシュデータの後にするために、ネクストポインタとして現在のMRUポインタが示すキャッシュデータのアドレスをセットし、ステップS55でMRUポインタとして今回接続すべきキャッシュデータのアドレスをセットして処理を終了する。

30

【0048】

ステップS52でMRUポインタが存在しないと判定された時には、キャッシュデータが存在していないことになり、ステップS56で今回接続すべきキャッシュデータの前後にキャッシュデータがない状態を示すためにネクストポインタ、プリビアスポインタそれぞれにNULL（終端）をセットし、ステップS57でMRU、LRUポインタのそれぞれに今回接続すべきキャッシュデータのアドレスをセットして処理を終了する。

【0049】

図11はRAIDキューからのデータの切断処理のフローチャートである。同図において処理が開始されると、ステップS61で今回切断すべきキャッシュデータの前後のキャッシュデータがリンクによって接続され、ステップS62でMRUポインタが獲得され、ステップS63でMRUポインタが指すキャッシュデータが今回切断すべきキャッシュデータであるか否かが判定され、そのキャッシュデータである場合にはMRUポインタが今回切断すべきキャッシュデータの後のキャッシュデータを指すように修正された後に、切断すべきキャッシュデータでない場合には直ちに、ステップS65でLRUポインタが指すキャッシュデータが切断すべきキャッシュデータであるか否かが判定され、そうである場合にはステップS66でLRUポインタが今回切断すべきキャッシュデータの前のキャッシュデータを指すように更新された後に、切断すべきキャッシュデータでない場合には

40

50

直ちに処理を終了する。

【0050】

図12は本実施形態におけるライトバックスケジュール処理の詳細フローチャートである。同図においてまずステップS71からステップS73で、通常のLRU法によるライトバックスケジュール処理が実行される。この処理ではまずステップS71でLRUキュー、すなわち図14の形式のキューの最後尾から最も使用頻度の低いデータが検索され、ステップS72でそのデータがライトバック対象のデータとして獲得され、ステップS73でそのデータに対するライトバック実行処理が行われる。この処理は図7で説明したものである。

【0051】

続いてステップS74からステップS78でRAID装置単位のライトバックスケジュール処理が実行される。まずステップS74でライトバック数の少ないグループが選択され、ステップS75でライトバック数がRAID装置毎にチェックされ、ライトバック数が少ないRAID装置がない場合には直ちに処理を終了する。ライトバック数が少ないRAID装置がある場合にはステップS76で空きキャッシュメモリ、すなわち図2のRAID制御装置のキャッシュメモリ16の中でライトバック数が少ないRAID装置12に割当てられている領域の中で空き領域があるか否かが判定され、空き領域がある場合にはライトバックをそれ以上行う必要がないものとして処理を終了する。

【0052】

空きキャッシュメモリがない場合には、ステップS77で該当するRAID装置にライトすべきデータがRAIDデータキューの最後尾から検索され、ステップS78でそのデータがライトバック対象のデータとして獲得され、ステップS79でライトバック実行処理が行われて処理を終了する。

【0053】

以上において本発明のRAID制御装置および制御方法についてその詳細を説明したが、このRAID制御装置は当然一般的なコンピュータシステムを基本として構成することが可能である。図13はそのようなコンピュータシステム、すなわちハードウェア環境の構成ブロック図である。

【0054】

図13においてコンピュータシステムは中央処理装置(CPU)20、リードオンメモリ(ROM)21、ランダムアクセスメモリ(RAM)22、通信インタフェース23、記憶装置24、入出力装置25、可搬型記憶媒体の読取り装置26、およびこれらの全てが接続されたバス27によって構成されている。

【0055】

記憶装置24としてはハードディスク、磁気ディスクなど様々な形式の記憶装置を使用することができ、このような記憶装置24、またはROM21に図3~図5、図7、図8、図10~図12などのフローチャートに示されたプログラムや、本発明の特許請求の範囲の請求項9のプログラムなどが格納され、そのようなプログラムがCPU20によって実行されることにより、本実施形態における複数のRAID装置へのキャッシュメモリの割当て、ライトバック実行数によるRAID装置のグループ化などが可能となる。

【0056】

このようなプログラムは、プログラム提供者28からネットワーク29、および通信インタフェース23を介して、例えば記憶装置24に格納されることも、また市販され、流通している可搬型記憶媒体30に格納され、読取り装置26にセットされて、CPU20によって実行されることも可能である。可搬型記憶媒体30としてはCD-ROM、フレキシブルディスク、光ディスク、光磁気ディスク、DVDなど様々な形式の記憶媒体を使用することができ、このような記憶媒体に格納されたプログラムが読取り装置26によって読取られることにより、本実施形態におけるすべてのRAID装置に対するまんべんないライトバックのスケジュールが可能となる。

【0057】

10

20

30

40

50

(付記 1)

複数の記憶装置を制御する記憶制御装置において、

該記憶制御装置内のキャッシュメモリに格納されているデータの前記複数の記憶装置へのライトバックをLRU法によって行うLRUライトバック手段と、

該LRUライトバック手段によって実行されているライトバックの数が少ない記憶装置を選択して、該選択した記憶装置に対するデータのライトバックを行うライトバックスケジュール処理手段とを備えることを特徴とする記憶制御装置。

【0058】

(付記 2)

前記記憶制御装置において、

前記複数の記憶装置を、該記憶装置が実行しているライトバックの数によって複数のグループに分割した結果を記憶するライトバックグループ管理手段をさらに備え、

前記ライトバックスケジュール処理手段が、ライトバックグループ管理手段の記憶内容に従ってライトバックの数が少ないグループ内の記憶装置を選択することを特徴とする付記1記載の記憶制御装置。

【0059】

(付記 3)

前記記憶制御装置において、

前記複数の記憶装置毎に実行されているライトバックの数を記憶する記憶装置管理手段をさらに備え、

前記ライトバックスケジュール処理手段が、該記憶装置管理手段の記憶内容に従ってライトバックの数が少ない記憶装置を選択することを特徴とする付記1記載の記憶制御装置。

【0060】

(付記 4)

前記記憶制御装置において、

該記憶制御装置内のキャッシュメモリの容量を分割して、前記複数の記憶装置に割り当てるキャッシュメモリ割り当て手段と、

該割り当てられた容量の中で該キャッシュメモリにデータが格納されていない未使用領域の容量を記憶装置毎に記憶する記憶装置管理手段とをさらに備え、

前記ライトバックスケジュール処理手段が、該記憶装置管理手段の記憶内容に従って該未使用領域の少ない記憶装置を選択することを特徴とする付記1記載の記憶制御装置。

【0061】

(付記 5)

前記記憶制御装置において、

前記記憶装置に割り当てられたキャッシュメモリの容量の内部に格納されているデータのキューを記憶装置毎に管理するキャッシュメモリ管理手段をさらに備え、

前記ライトバックスケジュール処理手段が、選択した記憶装置のキューにおけるLRU位置のデータのライトバックを行うことを特徴とする付記4記載の記憶制御装置。

【0062】

(付記 6)

自装置内のキャッシュメモリに格納されているデータの複数の外部の記憶装置へのライトバックをLRU法によって行うLRUライトバック手段と、

該LRUライトバック手段によって実行されているライトバックの数が少ない記憶装置を選択し、該選択した記憶装置に対するデータのライトバックを行うライトバックスケジュール処理手段とを備える記憶制御装置と、

該記憶制御装置によって制御される複数の記憶装置とによって構成されることを特徴とするデータ記憶システム。

【0063】

(付記 7)

10

20

30

40

50

前記記憶装置が複数のディスクを備える R A I D 装置であり、

前記記憶制御装置が該複数の R A I D 装置を制御する R A I D 制御装置であることを特徴とする付記 6 記載のデータ記憶システム。

【 0 0 6 4 】

(付記 8)

前記記憶制御装置が、前記複数の記憶装置を、該記憶装置が実行しているライトバックの数によって複数のグループに分割した結果を記憶するライトバックグループ管理手段をさらに備え、

前記ライトバックスケジュール処理手段が、ライトバックグループ管理手段の記憶内容に従ってライトバックの数が少ないグループ内の記憶装置を選択することを特徴とする付記 6 記載のデータ記憶システム。

10

【 0 0 6 5 】

(付記 9)

前記記憶制御装置が、前記複数の記憶装置毎に実行されているライトバックの数を記憶する記憶装置管理手段をさらに備え、

前記ライトバックスケジュール処理手段が、該記憶装置管理手段の記憶内容に従ってライトバックの数が少ない記憶装置を選択することを特徴とする付記 6 記載のデータ記憶システム。

【 0 0 6 6 】

(付記 1 0)

前記記憶制御装置が、該記憶制御装置内のキャッシュメモリの容量を分割して、前記複数の記憶装置に割当てられたキャッシュメモリ割当手段と、

該割当てられた容量の中で該キャッシュメモリにデータが格納されていない未使用領域の容量を記憶装置毎に記憶する記憶装置管理手段とをさらに備え、

前記ライトバックスケジュール処理手段が、該記憶装置管理手段の記憶内容に従って該未使用領域の少ない記憶装置を選択することを特徴とする付記 6 記載のデータ記憶システム。

20

【 0 0 6 7 】

(付記 1 1)

前記記憶制御装置が、前記記憶装置に割当てられたキャッシュメモリの容量の内部に格納されているデータのキューを記憶装置毎に管理するキャッシュメモリ管理手段をさらに備え、

前記ライトバックスケジュール処理手段が、選択した記憶装置のキューにおける L R U 位置のデータのライトバックを行うことを特徴とする付記 1 0 記載のデータ記憶システム。

30

【 0 0 6 8 】

(付記 1 2)

複数の記憶装置を制御する記憶制御方法において、

キャッシュメモリに格納されているデータの前記複数の記憶装置側へのライトバックを L R U 法によって実行し、

該 L R U 法によって実行されているライトバックの数が少ない記憶装置を選択して、該選択した記憶装置に対するデータのライトバックを実行することを特徴とする記憶制御方法。

40

【 0 0 6 9 】

(付記 1 3)

前記記憶制御方法において、

前記複数の記憶装置を、該記憶装置が実行しているライトバックの数によって複数のグループに分割し、

前記記憶装置の選択において、該記憶内容に従ってライトバックの数が少ないグループ内の記憶装置を選択することを特徴とする付記 1 2 記載の記憶制御方法。

50

【 0 0 7 0 】

(付記 1 4)

前記記憶制御方法において、
前記複数の記憶装置毎に実行されているライトバックの数を記憶し、
前記記憶装置の選択において、該記憶内容に従ってライトバックの数が少ない記憶装置を選択することを特徴とする付記 1 2 記載の記憶制御方法。

【 0 0 7 1 】

(付記 1 5)

前記記憶制御方法において、
前記記憶制御装置内のキャッシュメモリの容量を分割して、前記複数の記憶装置に割当
て、
該割当てられた容量の中で該キャッシュメモリにデータが格納されていない未使用領域
の容量を記憶装置毎に記憶し、
前記記憶装置の選択において、該記憶の内容に従って該未使用領域の少ない記憶装置を
選択することを特徴とする付記 1 2 記載の記憶制御方法。

【 0 0 7 2 】

(付記 1 6)

前記記憶制御方法において、
前記記憶装置に割当てられたキャッシュメモリの容量の内部に格納されているデータの
キューを記憶装置毎に管理し、
前記選択した記憶装置のキューにおける L R U 位置のデータのライトバックを行うこと
を特徴とする付記 1 5 記載の記憶制御方法。

【 0 0 7 3 】

(付記 1 7)

複数の記憶装置を制御する計算機によって使用されるプログラムであって、
キャッシュメモリに格納されているデータの前記複数の記憶装置側へのライトバックを
L R U 法によって実行する手順と、
該 L R U 法によって実行されているライトバックの数が少ない記憶装置を選択して、該
選択した記憶装置に対するデータのライトバックを実行する手順とを計算機に実行させる
ためのプログラム。

【 0 0 7 4 】

(付記 1 8)

前記記憶装置の選択において、
前記複数の記憶装置を、該記憶装置が実行しているライトバックの数によって複数のグ
ループに分割した結果を記憶するメモリの記憶内容に従ってライトバックの数が少ないグ
ループ内の記憶装置を選択することを特徴とする付記 1 7 記載のプログラム。

【 0 0 7 5 】

(付記 1 9)

前記記憶装置の選択において、
前記複数の記憶装置毎に実行されているライトバックの数を記憶するメモリの記憶内容
に従ってライトバックの数が少ない記憶装置を選択することを特徴とする付記 1 7 記載の
プログラム。

【 0 0 7 6 】

(付記 2 0)

前記記憶装置の選択において、
前記記憶制御装置内のキャッシュメモリの容量が分割され、前記複数の記憶装置に割当
てられた容量の中で該キャッシュメモリにデータが格納されていない未使用領域の容量を
記憶装置毎に記憶するメモリの記憶内容に従って該未使用領域の少ない記憶装置を選
択することを特徴とする付記 1 7 記載のプログラム。

【 0 0 7 7 】

(付記 2 1)

前記記憶装置に割当てられたキャッシュメモリの容量の内部に格納されているデータのキューが記憶装置毎に管理され、

前記選択した記憶装置のキューにおけるLRU位置のデータのライトバックを行うことを特徴とする付記 2 0 記載のプログラム。

【0078】

(付記 2 2)

複数の記憶装置を制御する計算機によって使用される記憶媒体であって、

キャッシュメモリに格納されているデータの前記複数の記憶装置側へのライトバックをLRU法によって実行するステップと、

該LRU法によって実行されているライトバックの数が少ない記憶装置を選択して、該選択した記憶装置に対するデータのライトバックを実行するステップとを計算機に実行させるプログラムを格納した計算機読出し可能可搬型記憶媒体。

10

【0079】

(付記 2 3)

前記記憶装置の選択において、

前記複数の記憶装置を、該記憶装置が実行しているライトバックの数によって複数のグループに分割した結果を記憶するメモリの記憶内容に従ってライトバックの数が少ないグループ内の記憶装置を選択することを特徴とする付記 2 2 記載の計算機読出し可能可搬型記憶媒体。

20

【0080】

(付記 2 4)

前記記憶装置の選択において、

前記複数の記憶装置毎に実行されているライトバックの数を記憶するメモリの記憶内容に従ってライトバックの数が少ない記憶装置を選択することを特徴とする付記 2 2 記載の計算機読出し可能可搬型記憶媒体。

【0081】

(付記 2 5)

前記記憶装置の選択において、

前記記憶制御装置内のキャッシュメモリの容量が分割され、前記複数の記憶装置に割当てられた容量の中で該キャッシュメモリにデータが格納されていない未使用領域の容量を記憶装置毎に記憶するメモリの記憶内容に従って該未使用領域の少ない記憶装置を選択することを特徴とする付記 2 2 記載の計算機読出し可能可搬型記憶媒体。

30

【0082】

(付記 2 6)

前記記憶装置に割当てられたキャッシュメモリの容量の内部に格納されているデータのキューが記憶装置毎に管理され、

前記選択した記憶装置のキューにおけるLRU位置のデータのライトバックを行うことを特徴とする付記 2 5 記載の計算機読出し可能可搬型記憶媒体。

【図面の簡単な説明】

40

【0083】

【図 1】本発明の記憶制御装置の原理構成ブロック図である。

【図 2】本実施形態におけるRAIDデータ記憶システムの全体構成ブロック図である。

【図 3】キャッシュメモリ分割処理のフローチャートである。

【図 4】キャッシュメモリ使用時処理のフローチャートである。

【図 5】キャッシュメモリ解放時処理のフローチャートである。

【図 6】RAID装置のグループ化の説明図である。

【図 7】ライトバック実行処理のフローチャートである。

【図 8】ライトバック完了処理のフローチャートである。

【図 9】キャッシュデータのキュー（リンク）の説明図である。

50

【図10】RAIDキューへの接続処理のフローチャートである。

【図11】RAIDキューからの切断処理のフローチャートである。

【図12】ライトバック処理の全体フローチャートである。

【図13】本発明を実現するためのプログラムのコンピュータへのローディングを説明する図である。

【図14】LRU法によるキャッシュデータ管理の従来例の説明図である。

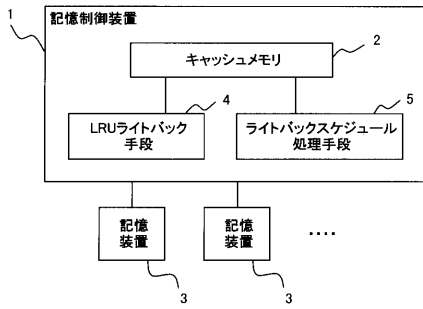
【符号の説明】

【0084】

1	記憶制御装置	
2	キャッシュメモリ	10
3	記憶装置	
4	LRUライトバック手段	
5	ライトバックスケジュール処理手段	
10	RAID制御装置	
11	ホストコンピュータ(サーバ)	
12	RAID装置	
13	チャンネルアダプタ	
14	デバイスアダプタ	
15	中央制御部	
16	キャッシュメモリ	20
17	キャッシュメモリ管理テーブル	
18	RAID装置毎の管理テーブル	
19	ライトバックグループ管理テーブル	
20	CPU	
21	リード・オンリ・メモリ(ROM)	
22	ランダム・アクセス・メモリ(RAM)	
23	通信インタフェース	
24	記憶装置	
25	入出力装置	
26	読取装置	30
27	バス	
28	プログラム提供者	
29	ネットワーク	
30	可搬型記憶媒体	

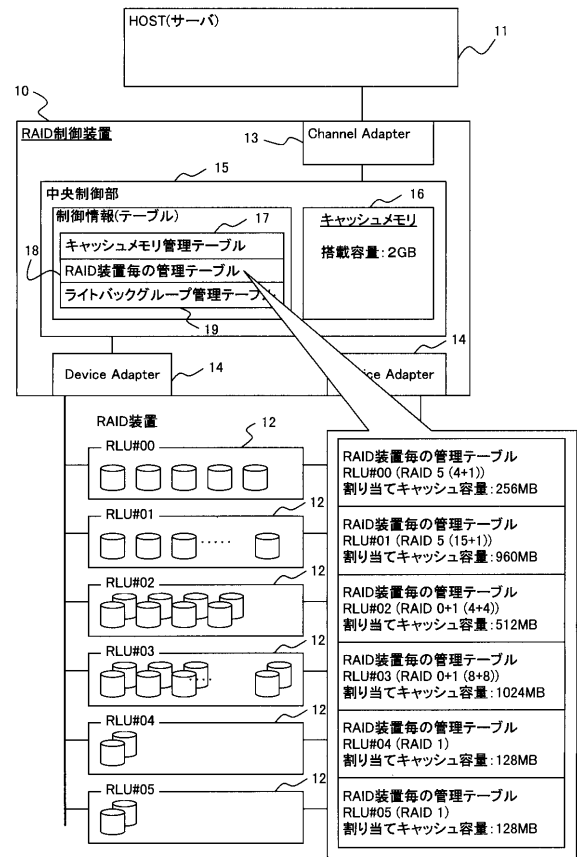
【 図 1 】

本発明の記憶制御装置の原理構成ブロック図



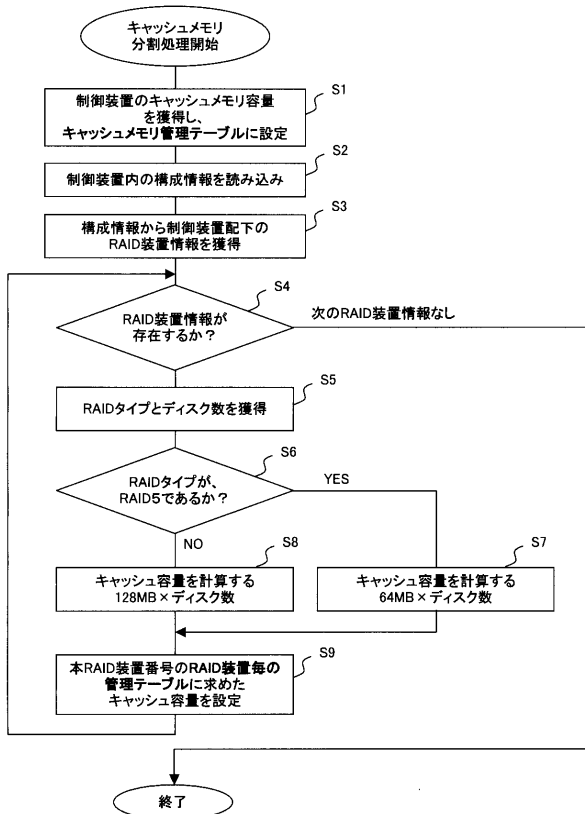
【 図 2 】

本実施形態におけるRAIDデータ記憶システムの全体構成ブロック図



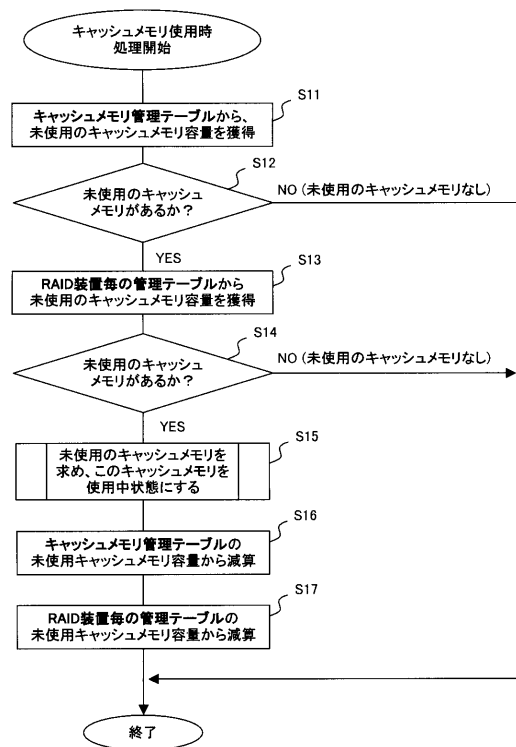
【 図 3 】

キャッシュメモリ分割処理のフローチャート



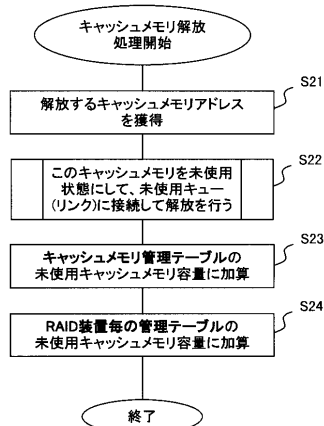
【 図 4 】

キャッシュメモリ使用時処理のフローチャート



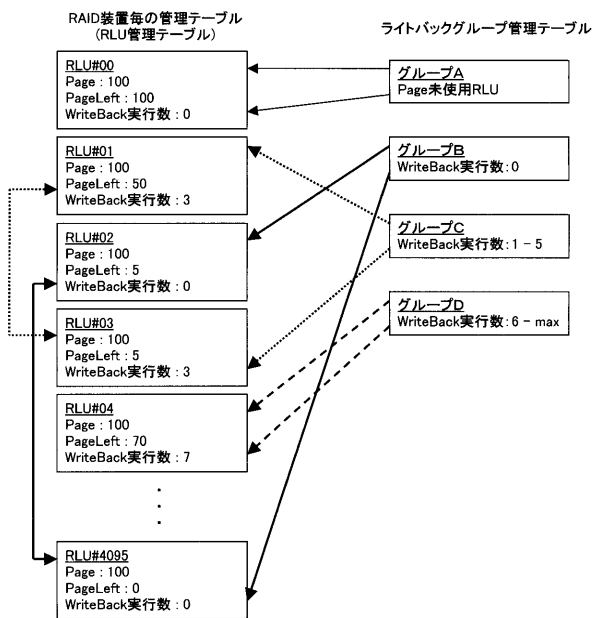
【 図 5 】

キャッシュメモリ解放時処理のフローチャート



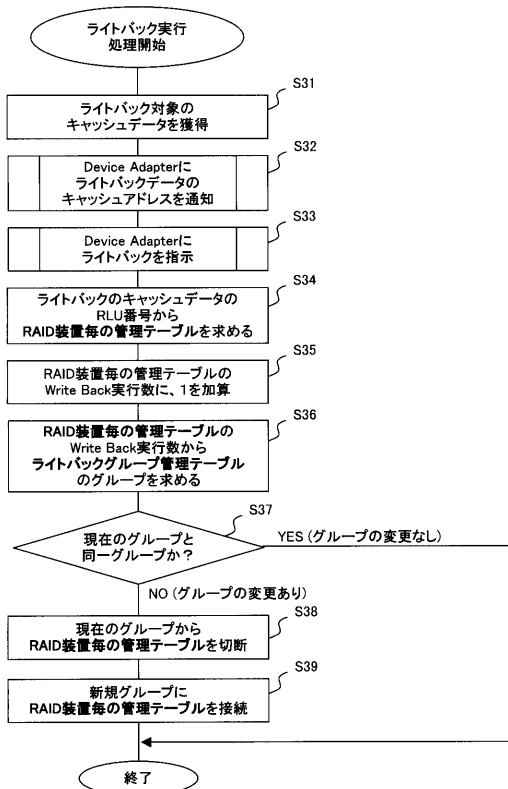
【 図 6 】

RAID装置のグループ化の説明図



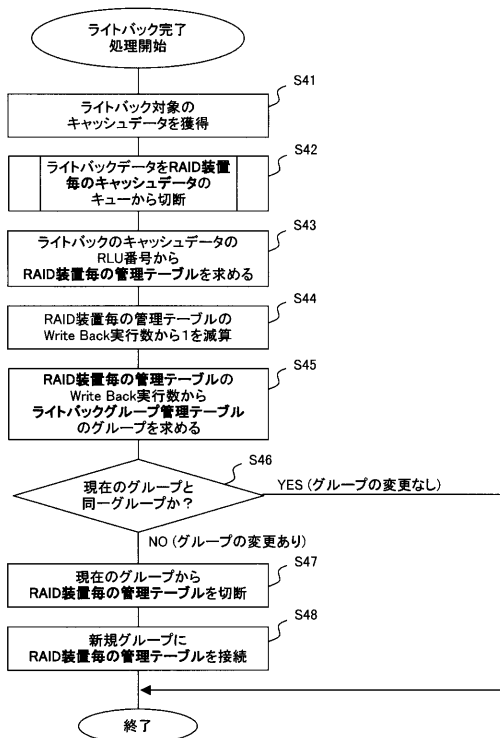
【 図 7 】

ライトバック実行処理のフローチャート



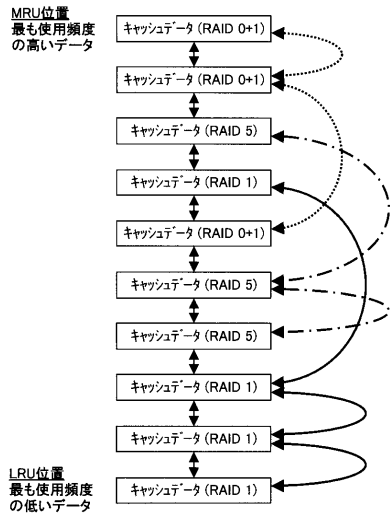
【 図 8 】

ライトバック完了処理のフローチャート



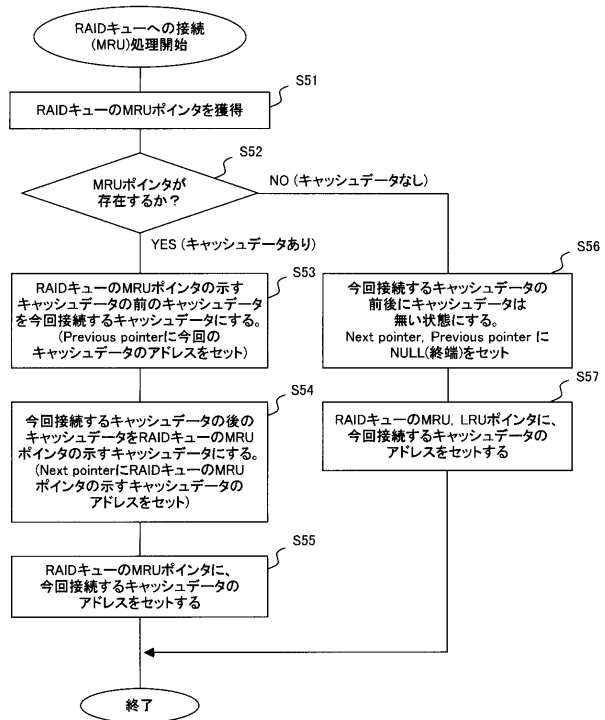
【 図 9 】

キャッシュデータのキュー(リンク)の説明図



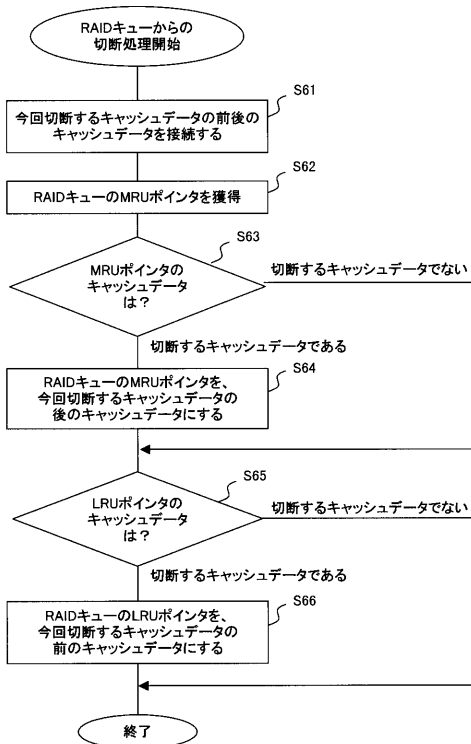
【 図 10 】

RAIDキューへの接続処理のフローチャート



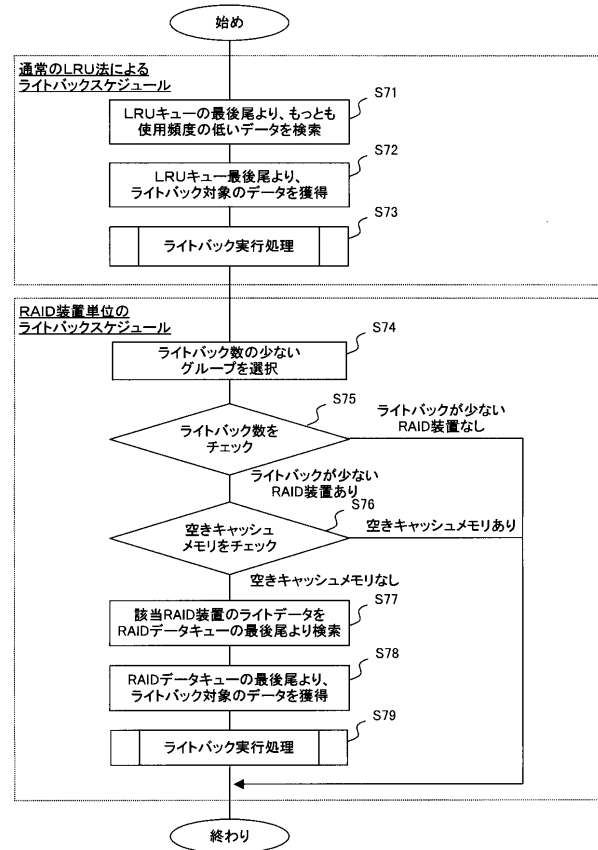
【 図 11 】

RAIDキューからの切断処理のフローチャート



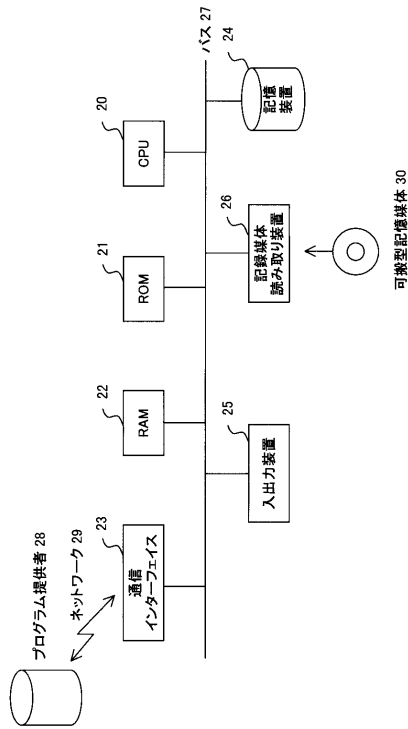
【 図 12 】

ライトバック処理の全体フローチャート



【 図 1 3 】

本発明を実現するためのプログラムのコンピュータへのローディングを説明する図



【 図 1 4 】

LRU法によるキャッシュデータ管理の従来例の説明図

