

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3541744号

(P3541744)

(45) 発行日 平成16年7月14日(2004.7.14)

(24) 登録日 平成16年4月9日(2004.4.9)

(51) Int. Cl.⁷

F I

G06F 3/06

G06F 3/06 301A

G06F 12/00

G06F 3/06 540

G06F 12/00 501B

請求項の数 10 (全 21 頁)

| | | | |
|-----------|----------------------------|-----------|--|
| (21) 出願番号 | 特願平11-242713 | (73) 特許権者 | 000005108 株式会社日立製作所 東京都千代田区神田駿河台四丁目6番地 |
| (22) 出願日 | 平成11年8月30日(1999.8.30) | (74) 代理人 | 100075096 弁理士 作田 康夫 |
| (65) 公開番号 | 特開2001-67187(P2001-67187A) | (72) 発明者 | 荒川 敬史 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 |
| (43) 公開日 | 平成13年3月16日(2001.3.16) | (72) 発明者 | 茂木 和彦 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 |
| 審査請求日 | 平成14年6月18日(2002.6.18) | | |

最終頁に続く

(54) 【発明の名称】 ストレージサブシステム及びその制御方法

(57) 【特許請求の範囲】

【請求項1】

複数の記憶装置と、前記記憶装置の使用状況情報を取得する手段と、前記計算機がリードライト対象とする論理記憶領域と前記記憶装置の第一の物理記憶領域との対応づけを行う手段とを有し、1台以上の計算機に接続するストレージサブシステムの制御方法であって、前記記憶装置は複数の組(クラス)に分類され、前記クラスは設定された属性を有し、前記ストレージサブシステムは、前記使用状況情報および前記クラス属性に基づき前記論理記憶領域に好適な再配置先のクラスを決定し、前記論理記憶領域の再配置先として利用可能な第二の物理記憶領域を前記クラス内から選択し、前記第一の物理記憶領域の内容を前記第二の前記物理記憶領域にコピーするとともに論理記憶領域の対応づけを前記第一の物理記憶領域から前記第二の物理記憶領域へ変更して再配置を行うことを特徴とする前記ストレージサブシステムの制御方法。

10

【請求項2】

請求項1に記載のストレージサブシステムの制御方法であって、ストレージサブシステムは、前記使用状況情報を蓄積し、設定された期間の前記使用状況情報に基づき、論理記憶領域の再配置先を決定し、設定された時間に再配置を行うことを特徴とするストレージサブシステムの制御方法。

【請求項3】

請求項1または2に記載のストレージサブシステムの制御方法であって、ストレージサブシステムは、使用状況情報として、記憶装置の単位時間当たりの使用時間(使用率)を用

20

い、各クラスは、属性として設定されたクラス間の性能順位と使用率上限値を有し、前記ストレージサブシステムは、クラスの使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択し、前記論理記憶領域の再配置先のクラスを前記順位の上位のクラスから、各クラスの使用率上限値を超えないように決定することを特徴とするストレージサブシステムの制御方法。

【請求項 4】

請求項 1 または 2 に記載のストレージサブシステムの制御方法であって、ストレージサブシステムは、使用状況情報として、記憶装置の単位時間当たりの使用時間（使用率）を用い、各クラスは、属性として設定されたクラス間の性能順位と使用率上限値を有し、前記ストレージサブシステムは、クラスの使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択し、前記論理記憶領域の再配置先として利用可能な物理記憶領域を同一クラス内の記憶装置から、前記クラスの使用率上限値を超えないように決定することを特徴とするストレージサブシステムの制御方法。

10

【請求項 5】

請求項 1 または 2 に記載のストレージサブシステムの制御方法であって、ストレージサブシステムは、使用状況情報として、記憶装置の単位時間当たりの使用時間（使用率）を用い、各クラスは属性として設定された対象アクセス種別と使用率上限値を有し、前記ストレージサブシステムは、クラスの使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択し、前記論理記憶領域に対するアクセス種別の分析結果に基づいて前記論理記憶領域の再配置先のクラスを前記対象アクセス種別のクラスから、各クラスの使用率

20

【請求項 6】

1 台以上の計算機に接続し、複数の記憶装置と、前記記憶装置の使用状況情報を取得する手段と、前記計算機がリードライト対象とする論理記憶領域と前記記憶装置の第一の物理記憶領域との対応づけを行う手段とを有するストレージサブシステムであって、前記複数のディスク装置をそれぞれ属性を有する複数の組（クラス）として管理する手段と、前記使用状況情報および前記クラス属性に基づき前記論理記憶領域に好適な再配置先のクラスを決定する手段と、前記論理記憶領域の再配置先として利用可能な第二の物理記憶領域を前記クラス内から選択する手段と、前記第一の物理記憶領域の内容を前記第二の前記物理記憶領域にコピーするとともに論理記憶領域の対応づけを前記第一の物理記憶領域から前記第二の物理記憶領域へ変更して再配置を行う手段とを有することを特徴とするストレージサブシステム。

30

【請求項 7】

請求項 6 に記載のストレージサブシステムであって、ストレージサブシステムは、前記使用状況情報を蓄積し、設定された期間の前記使用状況情報に基づき、論理記憶領域の再配置先を自動的に決定する手段と、設定された時間に再配置を行う手段とを有することを特徴とするストレージサブシステム。

【請求項 8】

請求項 6 または 7 に記載のストレージサブシステムであって、ストレージサブシステムは、使用状況情報として記憶装置の単位時間当たりの使用時間（使用率）を用いる手段を有し、前記ストレージサブシステムは、各クラスに属性として設定されている使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択する手段と、前記論理記憶領域の再配置先のクラスを各クラスに属性として設定されているクラス間の性能順位から、各クラスの使用率上限値を超えないように決定する手段とを有することを特徴とするストレージサブシステム。

40

【請求項 9】

請求項 6 または 7 に記載のストレージサブシステムであって、ストレージサブシステムは、使用状況情報として、記憶装置の単位時間当たりの使用時間（使用率）を用いる手段を有し、前記ストレージサブシステムは、属性として設定されたクラスの使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択する手段と、前記論理記憶領域に対

50

するアクセス種別を分析する手段と、対象アクセス種別を属性として設定されたクラスから、前記論理記憶領域の再配置先のクラスを前記分析結果に基づいて各クラスの使用率上限値を超えないように決定する手段とを有することを特徴とするストレージサブシステム。

【請求項10】

請求項6、7、8、または9に記載のストレージサブシステムであって、ストレージサブシステムは、複数のディスク装置を有するディスクアレイであり、前記ディスク装置の使用率を使用状況情報として用いる手段を有することを特徴とするストレージサブシステム。

【発明の詳細な説明】

10

【0001】

【発明の属する技術分野】

本発明は、複数の記憶装置を有するストレージサブシステム、およびその制御方法に関する。

【0002】

【従来の技術】

コンピュータシステムにおいて、高性能を実現する二次記憶システムの1つにディスクアレイシステムがある。ディスクアレイシステムは、複数のディスク装置をアレイ状に配置し、前記各ディスク装置に分割格納されるデータのリード/ライトを、前記各ディスク装置を並列に動作させることによって、高速に行うシステムである。ディスクアレイシステムに関する論文としては、D. A. Patterson, G. Gibson, and R. H. Kats, "A Case for Redundant Arrays of Inexpensive Disks (RAID)" (in Proc. ACM SIGMOD, pp. 109-116, June 1988)がある。この論文では、冗長性を付加したディスクアレイシステムに対し、その構成に応じてレベル1からレベル5の種別を与えている。これらの種別に加えて、冗長性無しのディスクアレイシステムをレベル0と呼ぶこともある。上記の各レベルは冗長性などにより実現するためのコストや性能特性などが異なるため、ディスクアレイシステムを構築するにあたって、複数のレベルのアレイ(ディスク装置の組)を混在させることも多い。ここでは、この組のことをパリティグループと呼ぶ。

20

30

【0003】

ディスク装置は、性能や容量などによりコストが異なり、ディスクアレイシステムを構築するにあたって最適なコストパフォーマンスを実現するために、やはり性能や容量の異なる複数種のディスク装置を用いることがある。

【0004】

ディスクアレイシステムに格納されるデータを上記のようにディスク装置に分散して配置するため、ディスクアレイシステムは、ディスクアレイシステムに接続するホストコンピュータがアクセスする論理記憶領域とディスク装置の記憶領域を示す物理記憶領域の対応づけ(アドレス変換)を行う。特開平9-274544号公報には、ホストコンピュータからの論理記憶領域に対するI/Oアクセスについての情報を取得する手段と、論理記憶領域の物理記憶領域への対応づけを変更して物理的再配置を行う手段により、格納されたデータの最適配置を実現するディスクアレイシステムが開示されている。

40

【0005】

【発明が解決しようとする課題】

特開平9-274544号公報に示されるような従来の技術における配置最適化の実行方法については以下の課題がある。

【0006】

再配置する論理記憶領域の選択および再配置先の物理記憶領域の選択にあたり、ディスクアレイシステムのユーザまたは保守員が、前記ディスクアレイシステムの構成や個々のディスク装置の特性や性能などの情報を確認して前記選択を行わなければならない、ユーザま

50

たは保守員による作業が煩雑となっていた。

【 0 0 0 7 】

また、ディスクアレイシステムが選択を自動的に行う場合においても、ユーザまたは保守員が前記個々のディスク装置の情報を確認して選択基準値を規定しなければならず、やはりユーザまたは保守員による作業が煩雑となっていた。特に、上記のように異種のレベルや異種のディスク装置の混在するディスクアレイシステムについては情報管理の煩雑さが増大する。

【 0 0 0 8 】

また、ディスクアレイシステムが選択のために行う I / O アクセス情報の参照は、ホストコンピュータおよびディスクアレイシステムを含むシステムで行われる処理のスケジュールの特性を考慮していなかった。一般にコンピュータシステムで行われる処理と処理に伴う I / O は、ユーザによって作成されたスケジュールに則って行われており、また処理および I / O の傾向は日毎、月毎、年毎などの周期性を示す場合も多く、一般にユーザは特定期間の処理および I / O に関心があると考えられる。

10

【 0 0 0 9 】

また上記従来技術において、再配置による性能チューニング方法については以下の課題がある。物理的再配置による性能チューニング方法は、ディスク装置、すなわち、物理記憶領域の使用状況に変更を加えるものであるが、従来技術においては、ホストコンピュータからの論理記憶領域に対する I / O アクセスについての情報を参照するため、再配置する論理記憶領域の選択および再配置先の物理記憶領域の選択にあたり、正しい選択が行えない可能性があった。

20

【 0 0 1 0 】

また、ホストコンピュータからのシーケンシャルアクセスとランダムアクセスが顕著に、同一のディスク装置に含まれる別々の物理記憶領域に対して行われる場合でも、シーケンシャルアクセスとランダムアクセスを異なるディスク装置に分離するために、再配置先のディスク装置を任意に特定して自動的再配置を行わせることはできなかった。一般に、ホストコンピュータからの処理要件として、データ長の小さいランダムアクセスには短時間での応答（高応答性能）が求められるが、同一ディスク装置にデータ長の大きいシーケンシャルアクセスが存在する場合、ランダムアクセスの応答時間はシーケンシャルアクセスの処理に阻害されて長くなり、応答性能は悪化してしまう。

30

【 0 0 1 1 】

本発明の第一の目的は、ディスクアレイシステムのユーザまたは保守員が再配置による配置最適化を行うための作業を簡便にすることにある。

【 0 0 1 2 】

本発明の第二の目的は、ホストコンピュータおよびディスクアレイシステムを含むシステムでの処理のスケジュールを考慮した再配置による配置最適化を可能にすることにある。

【 0 0 1 3 】

本発明の第三の目的は、再配置する論理記憶領域の選択および再配置先の物理記憶領域の選択にあたり、実際の記憶装置であるディスク装置の使用状況に基づく選択を行う、ディスクアレイシステムの制御方法およびディスクアレイシステムを提供することにある。

40

【 0 0 1 4 】

本発明の第四の目的は、ディスクアレイシステムにおける同一ディスク装置での顕著なシーケンシャルアクセスとランダムアクセスの混在に対し、再配置先のディスク装置を任意に特定して再配置によりシーケンシャルアクセスおよびランダムアクセスを異なるディスク装置に自動的に分離することができるようにすることにある。

【 0 0 1 5 】

【課題を解決するための手段】

上記の第一目的を実現するために、1台以上のホストコンピュータに接続するディスクアレイシステムは、配下の複数のディスク装置の使用状況情報を取得する手段と、ホストコンピュータがリード/ライト対象とする論理記憶領域とディスク装置の第一の物理記憶領

50

域との対応づけを行う手段とを有し、さらに、複数のディスク装置をそれぞれ属性を有する複数の組（クラス）として管理する手段と、使用状況情報およびクラス属性に基づき論理記憶領域に好適な再配置先のクラスを決定する手段と、論理記憶領域の再配置先として利用可能な第二の物理記憶領域をクラス内から選択する手段と、第一の物理記憶領域の内容を前記第二の前記物理記憶領域にコピーするとともに論理記憶領域の対応づけを第一の物理記憶領域から第二の物理記憶領域へ変更して再配置を行う手段を備える。

【0016】

また、上記第二の目的を実現するために、ディスクアレイシステムは、使用状況情報を蓄積し、設定された期間の使用状況情報に基づき、論理記憶領域の再配置先を決定する手段と、設定された時間に再配置を行う手段を備えることができる。

10

【0017】

また、上記第三の目的を実現するために、ディスクアレイシステムは、使用状況情報として、ディスク装置の単位時間当たりの使用時間（使用率）を用いる手段を備える。

【0018】

また、上記第四の目的を実現するために、ディスクアレイシステムは、各クラスに属性として設定された対象アクセス種別（シーケンシャル/ランダムアクセス種別）と使用率上限値を用いて、クラスの使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択し、論理記憶領域に対するアクセス種別の分析結果に基づいて論理記憶領域の再配置先のクラスを好適なアクセス種別のクラスから、各クラスの使用率上限値を超えないように決定する手段を備える。

20

【0019】

【発明の実施の形態】

以下、本発明の実施の形態を図1～図27を用いて説明する。

【0020】

< 第一の実施の形態 >

本実施の形態では、クラス600に基づく再配置の判断と、再配置判断および実行のスケジューリングについて説明する。

【0021】

図1は、本発明の第1の実施の形態における計算機システムの構成図である。

【0022】

本実施の形態における計算機システムは、ホスト100、ストレージサブシステム200、制御端末700を有してなる。

30

【0023】

ホスト100は、ストレージサブシステム200にI/Oバス800を介して接続し、ストレージサブシステム200に対しリードやライトのI/Oを行う。I/Oの際、ホスト100は、ストレージサブシステム200の記憶領域について論理領域を指定する。I/Oバス800の例としては、ESCON、SCSI、ファイバチャネルなどがある。

【0024】

ストレージサブシステム200は、制御部300および複数の記憶装置500を有する。制御部300は、リード/ライト処理310、使用状況情報取得処理311、再配置判断処理312、及び再配置実行処理313を行う。また、ストレージサブシステム200は、論理/物理対応情報400、クラス構成情報401、クラス属性情報402、論理領域使用状況情報403、物理領域使用状況情報404、再配置判断対象期間情報405、再配置実行時刻情報406、未使用領域情報407、及び再配置情報408を保持する。

40

【0025】

ホスト100、制御部300、および制御端末700は、ネットワーク900で接続される。ネットワーク900の例としては、FDDI、ファイバチャネルなどがある。

【0026】

ホスト100、制御部300、および制御端末700には、各々での処理を行うためのメモリ、CPUなど、計算機において一般に用いられる構成要素もそれぞれ存在するが、本

50

実施の形態の説明においては重要でないため、ここでは説明を省略する。

【0027】

ホスト100が、ストレージサブシステム200に対してリード/ライトを行う場合のリード/ライト処理310、および使用状況情報取得処理311について図2で説明する。

【0028】

リード/ライト処理310において、ホスト100は、ストレージサブシステム200の制御部300に対しリードまたはライトを論理領域を指定して要求する(ステップ1000)。要求を受領した制御部300は、論理/物理対応情報400を用いて論理領域に対応する物理領域を求め、すなわち論理領域のアドレス(論理アドレス)を物理領域のアドレス(物理アドレス)に変換する(ステップ1010)。続いて制御部300は、リードの場合、この物理アドレスの記憶装置500からデータを読み出してホスト100に転送し、ライトの場合は、ホスト100から転送されたデータを前記物理アドレスの記憶装置500に格納し(ステップ1020)、さらに後述の使用状況情報取得処理311を行う。リード/ライト要求およびデータ転送はI/Oバス800を介して行われる。

10

【0029】

論理/物理対応情報400の一例を図3に示す。論理アドレスはホスト100がリード/ライト処理310で用いる論理領域を示すアドレスである。物理アドレスは実際にデータが格納される記憶装置500上の領域を示すアドレスであり、記憶装置番号および記憶装置内アドレスからなる。記憶装置番号は個々の記憶装置500を示す。記憶装置内アドレスは記憶装置500内での記憶領域を示すアドレスである。

20

【0030】

次に、使用状況情報取得処理311において制御部300は、リード/ライト処理310においてリード/ライト対象となった論理領域についての論理領域使用状況情報403と、リード/ライト処理310で使用した物理領域についての物理領域使用状況情報404を更新する(ステップ1030、1040)。論理領域使用状況情報403および物理領域使用状況情報404は、例えば使用頻度、使用率、リード/ライトに関する属性など、各々の論理領域と物理領域の各日時の使用状況に関する情報である。論理領域使用状況情報403および物理領域使用状況情報404の具体的な例は、以降の実施の形態で説明する。

【0031】

次に、制御部300が行う再配置判断処理312について図4で説明する。

30

【0032】

記憶装置500は、ユーザによって、または初期状態として複数の組(クラス600)に分類されており、クラス600への分類はクラス構成情報401に設定されている。さらに、各クラス600は、ユーザによって、または初期条件として属性を設定されており、属性は、クラス属性情報402に設定されている。クラス属性情報402は、許容使用状況や好適な使用状況やクラス間優先順位などの属性に関する情報である。クラス構成情報401およびクラス属性情報402の具体的な例は、以降の実施の形態で説明する。再配置判断対象期間情報405には、ユーザによってまたは初期条件として再配置判断処理312の対象とする使用状況情報の期間と期間更新情報が設定されている。

40

【0033】

再配置判断対象期間情報405の一例を図5に示す。開始日時から終了日時までの期間が対象期間となる。期間更新情報は次の対象期間の設定条件であり、例えば毎週、毎日、X時間後などがありうる。制御部300は、対象期間の論理領域使用状況情報403および物理領域使用状況情報404を参照し(ステップ1100)、クラス属性情報402の各クラス600の許容使用状況などと比較して(ステップ1110)、物理的再配置を行うべき論理領域を選択する(ステップ1120)。

【0034】

さらに、制御部300は、クラス属性情報402の許容使用状況や好適な使用状況やクラス間優先順位などを参照して(ステップ1130)、論理領域の再配置先のクラス600

50

を選択し(ステップ1140)、さらに、クラス600に属する記憶装置500の中から論理領域の再配置先として未使用の物理領域を選択し(ステップ1150)、選択結果を再配置情報408に出力する(ステップ1160)。

【0035】

再配置情報408の一例を図6に示す。論理領域は、再配置する論理領域であり、再配置元物理領域は、論理領域に対応する現在の物理領域を示す記憶装置番号と記憶装置内アドレスであり、再配置先物理領域は、再配置先の物理領域を示す記憶装置番号と記憶装置内アドレスである。図6に示すように再配置の立案は一つ以上行われうる。さらに制御部300は、再配置判断対象期間情報405の期間更新情報を参照して、再配置判断対象期間情報405の対象期間を次回分に更新する(ステップ1170)。上記の処理において制

10

【0036】

未使用領域情報407の一例を図7に示す。記憶装置番号は個々の記憶装置500を示す。記憶装置内アドレスは記憶装置500内での領域を示すアドレスである。記憶装置番号および装置内アドレスは物理領域を示し、使用/未使用の項目は、物理領域の使用/未使用の区別を示す。制御部300は、通常、再配置判断処理312を対象期間以後、後述の再配置実行処理313以前に自動的に行う。

【0037】

次に、制御部300が行う再配置実行処理313について図8で説明する。

20

【0038】

再配置実行時刻情報406にはユーザによってまたは初期条件として再配置実行処理313を行う日時と日時更新情報が設定されている。

【0039】

再配置実行時刻情報406の一例を図9に示す。制御部300は、設定された日時に以下に説明する再配置実行処理313を自動的に実行する。日時更新情報は次回の再配置実行処理313を行う日時の設定条件であり、例えば毎週、毎日、X時間後などがありうる。制御部300は、再配置情報408に基づき再配置元物理領域に格納している内容を再配置先物理領域にコピーする(ステップ1200)。さらに、コピーが完了して再配置元物理領域の内容が全て再配置先物理領域に反映された時点で、制御部300は、論理/物理

30

【0040】

さらに、制御部300は、未使用物理領域470上の再配置先物理領域を使用とし、再配置元物理領域を未使用に変更する(ステップ1220)。さらに制御部300は、再配置実行時刻情報406の日時更新情報を参照して、再配置実行時刻情報406の日時を次回分に更新する(ステップ1230)。

【0041】

ユーザまたは保守員は、制御部300が上記の処理で用いている各情報を、制御端末700からネットワーク900を介して、またはホスト100からネットワーク900または

40

【0042】

上記の処理を行うことによって、取得した使用状況情報および設定されたクラス属性に基づいて、ストレージサブシステム200において論理領域の物理的再配置を自動的に行い、ストレージサブシステム200の最適化を行うことができる。さらに上記の再配置判断および実行の処理を繰り返して配置を修正していくことによって、使用状況の変動やその他の最適化誤差要因を吸収していくことができる。

【0043】

特に、上記の処理により、ユーザまたは保守員は再配置による最適化を簡便に行うことが

50

できる。ユーザまたは保守員は、記憶装置500をクラス600という単位で管理できるため、記憶装置500の性能や信頼性や特性などの属性を個々の前記記憶装置500について管理する必要はない。さらに、ユーザまたは保守員は、記憶装置500の個々の属性が等しくない組に対しても、必要に応じて同一の属性を持つクラス600を設定して、1つの管理単位として扱うことができる。ただし、1つの記憶装置500が1つのクラス600を構成すると見なして1つの記憶装置500を管理単位として上記の再配置の処理を行うことも可能である。

【0044】

また、ユーザまたは保守員は、ホスト100で行われる処理(ジョブ)の特徴やスケジュールを考慮して、上記の再配置を自動的に行うことができる。一般に、計算機システムで行われる処理と、この処理に伴うI/Oは、ユーザによって作成されたスケジュールに則って行われる。ユーザは、特に最適化の対象としたい処理を有する場合、処理の期間を特定することが可能であり、本実施の形態で説明した再配置の処理によって、ユーザは関心のある期間を指定して再配置判断の処理をストレージシステム200に行わせ、すなわち、前記期間の使用状況情報に基づいて上記の再配置による最適化を実現することができる。また、計算機システムで行われる処理およびI/Oの傾向は日毎、月毎、年毎などの周期性を示す場合も多い。特に、処理が定型業務に基づく処理である場合には、周期性が顕著となる。前述の場合と同様にユーザは、周期において特に最適化対象として関心のある期間を指定して再配置による最適化を行うことができる。また、再配置実行処理313では、ストレージシステム200内で格納内容のコピーを伴うが、ユーザはストレージシステム200があまり使用されていない時刻やホスト100で実行されている処理の要求処理性能が低い期間を再配置実行処理313の実行時刻として設定することで、ホスト100での要求処理性能が高い処理のストレージシステム200へのI/Oがコピーにより阻害されることを回避できる。

【0045】

なお、記憶装置500は、それぞれ異なる性能、信頼性、特性や属性を持っていてよく、特に具体的には、磁気ディスク装置、磁気テープ装置、半導体メモリ(キャッシュ)のように異なる記憶媒体であってもよい。また、上記の例では未使用領域情報407は物理領域に基づいて記述されているとしたが、未使用の物理領域に対応する論理領域(論理アドレス)に基づいて記述されていてもよい。

【0046】

< 第二の実施の形態 >

本実施の形態では、使用状況情報としてのディスク装置使用率の適用と、クラス600の上限値およびクラス600間の性能順位による再配置判断について説明する。

【0047】

図10は、本発明の第2の実施の形態における計算機システムの構成図である。

【0048】

本実施の形態の計算機システムは、ホスト100、ディスクアレイシステム201、制御端末700を有してなる。本実施の形態における計算機システムは、第1の実施の形態でのストレージサブシステム200をディスクアレイシステム201とし、記憶装置500をパリティグループ501としたものに相当する。

【0049】

ディスクアレイシステム201は、制御部300とディスク装置502を有する。制御部300は、第1の実施の形態での制御部300に相当する。ディスク装置502は、n台(nは2以上の整数)でRAID(ディスクアレイ)を構成しており、このn台のディスク装置502による組をパリティグループ501と呼ぶ。RAIDの性質として、1つのパリティグループ501に含まれるn台のディスク装置502は、n-1台のディスク装置502の格納内容から生成される冗長データが残りの1台に格納されるといった冗長性上の関係を持つ。またn台のディスク装置502は、冗長データを含めた格納内容が、並列動作性向上のためにn台のディスク装置502に分散格納されるなど、データ格納上の

10

20

30

40

50

関係を持つ。この関係から各パリティグループ501を動作上の1単位とみなすことができるが、冗長性や台数nなどにより実現するためのコストや性能特性などが異なるため、ディスクアレイシステム201を構成するにあたって、レベルや台数nの異なるアレイ(パリティグループ501)を混在させることも多く、またパリティグループ501を構成するディスク装置502についても、性能や容量などによりコストが異なるため、ディスクアレイシステム201を構成するにあたって最適なコストパフォーマンスを実現するために性能や容量の異なる複数種のディスク装置502を用いることもある。よって本実施の形態においてディスクアレイシステム201を構築する各パリティグループ501は性能、信頼性、特性などの属性が同一であるとは限らず、特に性能について差異があるとす

10

【0050】

本実施の形態における論理/物理対応情報400の一例を図11に示す。

【0051】

論理アドレスは、ホスト100がリード/ライト処理310で用いる論理領域を示すアドレスである。物理アドレスは実際にデータと前記冗長データが格納されるディスク装置502上の領域を示すアドレスであり、パリティグループ番号と各々のディスク装置番号およびディスク装置内アドレスからなる。パリティグループ番号は個々のパリティグループ501を示す。ディスク装置番号は個々のディスク装置502を示す。ディスク装置内アドレスはディスク装置502内での領域を示すアドレスである。制御部300は、RAIDの動作として、冗長データに関する情報を前記リード/ライト処理310などで用いて

20

【0052】

さらに第1の実施の形態と同様に、パリティグループ501は、ユーザによってまたは初期状態として複数の組(クラス600)に分類されており、クラス600への分類はクラス構成情報401に設定されている。クラス構成情報401の一例を図12に示す。

【0053】

クラス番号は各クラス600を示す番号である。パリティグループ数は各クラス600に属するパリティグループの数を示す。パリティグループ番号は各クラス600に属するパリティグループ番号501を示す。同様に、各クラス600の属性は、クラス属性情報4

30

【0054】

クラス番号は、各クラス600を示す番号である。使用率上限値は後述のディスク使用率の許容範囲を示す上限値であり、クラス600の属するパリティグループ501に適用する。クラス間性能順位は、クラス600間の性能順位(数字の小さいものが高性能とする)である。クラス間性能順位は各クラス600を構成するパリティグループ501の前述の性能差異に基づく。再配置実行上限値および固定については後述する。

【0055】

本実施の形態における使用状況情報取得処理311について図14で説明する。

40

【0056】

制御部300は、第1の実施の形態と同様に、リード/ライト処理310において使用したディスク装置502の使用時間を取得して単位時間当たりの使用時間(使用率)を求め、さらに、ディスク装置502が属するパリティグループ501について、使用率の平均を算出し(ステップ1300)、使用率平均を、リード/ライト対象となった論理領域についてのディスク装置使用率として論理領域使用状況情報403に記録する(ステップ1310)。また制御部300は、パリティグループ501に対応する全論理領域のディスク装置使用率の和を求め(ステップ1320)、パリティグループ501の使用率として物理領域使用状況情報404に記録する(ステップ1330)。

【0057】

50

本実施の形態における論理領域使用状況情報 4 0 3 および物理領域使用状況情報 4 0 4 の一例を図 1 5 および図 1 6 に示す。

【 0 0 5 8 】

日時はサンプリング間隔（一定期間）毎の日時を示し、論理アドレスは論理領域を示し、パリティグループ番号は個々のパリティグループを示し、論理領域のディスク装置使用率およびパリティグループ使用率はそれぞれ前記サンプリング間隔での平均使用率を示す。上記のようなディスク装置 5 0 2 の使用率はディスク装置 5 0 2 にかかる負荷を示す値であり、使用率が大きい場合は、ディスク装置 5 0 2 が性能ボトルネックとなっている可能性があるため、再配置処理で使用率を下げることによりディスクアレイシステム 2 0 1 の性能向上が期待できる。

10

【 0 0 5 9 】

次に、再配置判断処理 3 1 2 について図 1 7 で説明する。

【 0 0 6 0 】

制御部 3 0 0 は、各クラス 6 0 0 について、クラス 6 0 0 に属するパリティグループ 5 0 1 をクラス構成情報 4 0 1 から取得する（ステップ 1 3 0 0）。続いて、制御部 3 0 0 は、第 1 の実施の形態と同様の再配置判断対象期間情報 4 0 5 を参照して対象期間を取得し、さらにパリティグループ 5 0 1 について、対象期間の物理領域使用状況情報 4 0 4 のパリティグループ使用率を取得し集計する（ステップ 1 3 2 0）。続いて、制御部 3 0 0 は、クラス属性情報 4 0 2 を参照してクラス 6 0 0 の使用率上限値を取得する（ステップ 1 3 3 0）。制御部 3 0 0 は、パリティグループ使用率とクラス上限値を比較し、パリティ

20

【 0 0 6 1 】

続いて、制御部 3 0 0 は、対象期間の論理領域使用状況情報 4 0 3 を参照して、再配置が必要と判断したパリティグループ 5 0 1 の各物理領域に対応する論理領域のディスク装置使用率を取得し集計して（ステップ 1 3 5 0）、ディスク装置使用率の大きいものから、再配置する論理領域として選択する（ステップ 1 3 6 0）。論理領域の選択は、パリティグループ 5 0 1 の使用率から選択した論理領域のディスク使用率を減算していき、クラス 6 0 0 の使用率上限値以下になるまで行う（1 3 7 0）。ディスク装置使用率の大きい論

30

【 0 0 6 2 】

制御部 3 0 0 は、選択された論理領域についての再配置先となる物理領域を探す。制御部 3 0 0 は、クラス属性情報 4 0 2 を参照し、パリティグループ 5 0 1 の属するクラス 6 0 0 より性能順位が高位のクラス 6 0 0（高性能クラス）に注目し、クラス構成情報 4 0 1 および第 1 の実施の形態と同様の未使用領域情報 4 0 7 を参照して高性能クラスに属するパリティグループ 5 0 1 の未使用物理領域を取得する（ステップ 1 3 8 0）。

40

【 0 0 6 3 】

さらに、制御部 3 0 0 は、各未使用物理領域について、再配置先とした場合のパリティグループ使用率の予測値を求め（ステップ 1 3 9 0）、未使用物理領域の中から、再配置先とした場合に高性能クラスに設定されている上限値を超えないと予測できる未使用物理領域を、再配置先の物理領域として選択し（ステップ 1 4 0 0）、選択結果を第 1 の実施の形態と同様に、再配置情報 4 0 8 に出力する（ステップ 1 4 1 0）。選択した全ての論理領域について再配置先の物理領域を選択し終わったら処理を終了する（ステップ 1 4 2 0）。

【 0 0 6 4 】

本実施の形態において、制御部 3 0 0 は、第 1 の実施の形態に加えてパリティグループ情

50

報 4 0 9 を保持し、パリティグループ情報 4 0 9、論理領域使用状況情報 4 0 3、及び物理領域使用状況情報 4 0 4 から使用率予測値を算出する。

【 0 0 6 5 】

パリティグループ情報 4 0 9 の一例を図 1 8 に示す。パリティグループ番号は個々のパリティグループ 5 0 1 を示す番号である。RAID 構成はパリティグループ 5 0 1 が構成する RAID のレベルやディスク台数や冗長度構成を示す。ディスク装置性能はパリティグループ 5 0 1 を構成するディスク装置 5 0 2 の性能特性を示す。固定については後述する。上記の処理においてディスク装置使用率の大きい論理領域の再配置先を高性能クラスのパリティグループ 5 0 1 とすることで、同一負荷に対するディスク装置使用時間を短縮でき、論理領域の再配置後のディスク装置使用率を抑制できる。

10

【 0 0 6 6 】

再配置実行処理 3 1 3 は、第 1 の実施の形態と同様に行われるが、図 1 9 に示すように、制御部 3 0 0 は、再配置のためのコピーを行う前にクラス属性情報 4 0 2 を参照し、再配置元および再配置先のクラス 6 0 0 について、ユーザによってまたは初期条件として設定された再配置実行上限値を取得する（ステップ 1 5 0 0）。さらに物理領域使用状況情報 4 0 4 を参照して、再配置元および再配置先のパリティグループ 5 0 1 の直近のパリティグループ使用率を取得し（ステップ 1 5 1 0）、比較の結果少なくとも一方のクラス 6 0 0 においてパリティグループ使用率が再配置実行上限値を超えていた場合は（ステップ 1 5 2 0、1 5 3 0）、再配置実行処理 3 1 3 を中止または延期する（ステップ 1 5 4 0）。

20

【 0 0 6 7 】

上記処理によりユーザは、パリティグループ 5 0 1 の使用率が大きくすなわち負荷が高い場合に前記コピーによりさらに負荷が生じることを回避することができ、また回避のため上限値をクラス 6 0 0 毎に任意に設定することができる。

【 0 0 6 8 】

上記のように処理することによって、ディスク装置 5 0 2 の使用状況に基づいて物理的に再配置する論理領域の選択、および再配置先の物理領域の選択を、クラス構成および属性に基づいて行い、再配置によりディスク装置 5 0 2 の負荷を分散して、各クラス 6 0 0 に設定されている使用率上限値を、クラス 6 0 0 に属するパリティグループ 5 0 1 の使用率が超えない配置を実現することができる。さらに再配置判断および実行の処理を繰り返して配置を修正していくことによって、使用状況の変動や予測誤差を吸収していくことができる。

30

【 0 0 6 9 】

再配置判断処理 3 1 2 において、制御部 3 0 0 は、対象期間の物理領域使用状況情報 4 0 4 のパリティグループ使用率や、論理領域使用状況情報 4 0 3 の論理領域のディスク装置使用率を参照して集計し、判断に用いるとしたが、例えば、対象期間の全ての値の平均を用いる代わりに、対象期間中の上位 m 個の値を用いる方法も考えられ、また上位 m 番目の値を用いる方法も考えられる（m は 1 以上の整数）。これらの方法をユーザが選択できるようにすることで、ユーザは使用状況の特徴的な部分のみを選択して用い、再配置判断処理 3 1 2 を行わせることができる。

40

【 0 0 7 0 】

上記の再配置判断処理 3 1 2 において、制御部 3 0 0 は、ディスクアレイシステム 2 0 1 の全てのクラス 6 0 0 について、論理領域の再配置の必要なパリティグループ 5 0 1 の検出を行うとしたが、前記検出の前に制御部 3 0 0 がクラス属性情報 4 0 2 を参照し、固定属性が設定されているクラス 6 0 0 については、検出の対象外としてもよい。また同様に、制御部 3 0 0 がパリティグループ情報 4 0 9 を参照し、固定属性が設定されているパリティグループ 5 0 1 については検出の対象外としてもよい。また、再配置判断処理 3 1 2 において、制御部 3 0 0 は、高性能クラスに属するパリティグループ 5 0 1 の未使用物理領域から再配置先の物理領域を選択するとしたが、固定属性が設定されているクラス 6 0 0 については対象外として、さらに性能順位が高位のクラス 6 0 0 を高性能クラスとして

50

扱うようにしてもよい。また固定属性が設定されているパリティグループ501については対象外としてもよい。上記のように固定属性が設定されているクラス600またはパリティグループ501を扱うことによって、ユーザは上記の自動的な再配置処理において物理的な再配置の影響を生じさせたくないクラス600またはパリティグループ501を設定し、再配置の対象外とすることができる。

【0071】

< 第三の実施の形態 >

本実施の形態では、同一クラス600内での再配置判断について説明する。本実施の形態での計算機システムは、第2の実施の形態と同様である。ただし、本実施の形態では1つのクラス600に複数のパリティグループ501が属する。本実施の形態での処理は、再配置判断処理312を除いては第2の実施の形態と同様である。また、再配置判断処理312についても、再配置する論理領域の選択(ステップ1600)は、第2の実施の形態と同様である。

10

【0072】

本実施の形態での再配置判断処理312における、再配置先の物理領域の選択について図20で説明する。

【0073】

第2の実施の形態では再配置先の物理領域を再配置元の物理領域の属するクラス600より性能順位が高位のクラス600から選択するが、本実施の形態では同一クラス600の再配置元以外のパリティグループ501から選択する。制御部300は、クラス構成情報401と未使用領域情報407を参照して、同一クラス600に属する再配置元以外のパリティグループ501の未使用物理領域を取得する(ステップ1610)。制御部300は、各未使用物理領域について、再配置先とした場合のパリティグループ使用率の予測値を求め(ステップ1620)、未使用物理領域の中から、再配置先とした場合に同一クラス600に設定されている上限値を超えないと予測できる未使用物理領域を、再配置先の物理領域として選択し(ステップ1630)、選択結果を第2の実施の形態同様に、再配置情報408に出力する(ステップ1640)。再配置する全ての論理領域について再配置先の物理領域を選択し終えたら処理を終了する(ステップ1650)。

20

【0074】

上記の処理により、同一クラス600内においてディスク装置502の負荷を分散することができる。上記の処理方法は例えばディスクアレイシステム201のパリティグループ501が全て1つのクラス600(単一クラス)に属する構成に適用することができる。また、例えば、第2の実施の形態で説明した処理方法と組み合わせた場合に、再配置先の未使用物理領域の選択において、再配置元のクラス600より性能順位が高位のクラス600に適当な未使用物理領域が得られなかった場合や、性能順位が最上位のクラス600での処理に適用できる。第2の実施の形態で説明した処理方法と組み合わせた場合は、第2の実施の形態での処理方法と本実施の形態での処理方法とが各クラス600について異なる使用率上限値を用いてもよく、すなわち、そのためにクラス属性情報402が各クラス600について二種類の使用率上限値または差分を有してもよい。

30

【0075】

< 第四の実施の形態 >

本実施の形態では、第2の実施の形態での再配置判断処理312において、再配置元のクラス600より性能順位が高位のクラス600(高性能クラス)に再配置先の未使用物理領域が見つからなかった場合に、再配置先を得るために先立って行われる、性能順位がより低位のクラス600(低性能クラス)への高性能クラスからの再配置の処理について説明する。

40

【0076】

本実施の形態での計算機システムは、第2の実施の形態と同様である。本実施の形態における再配置判断処理312について図21で説明する。

【0077】

50

制御部300は、高性能クラスに属するパリティグループ501をクラス構成情報401から取得する(ステップ1700)。続いて制御部300は、第1の実施の形態と同様の再配置判断対象期間情報405を参照して対象期間を取得し(ステップ1710)、対象期間の論理領域使用状況情報403を参照して、パリティグループ501の各物理領域に対応する論理領域のディスク装置使用率を取得し(ステップ1720)、ディスク装置使用率の小さいものから、低性能クラスへ再配置する論理領域として選択する(ステップ1730)。このとき論理領域の選択は必要なだけ行われる(ステップ1740)。

【0078】

続いて制御部300は、選択された論理領域についての再配置先となる物理領域を、低性能クラスに属するパリティグループ501から選択するが、再配置先の物理領域選択の処理は、第2の実施の形態での処理説明において再配置先としている高性能クラスを低性能クラスと読み替えれば、第2の実施の形態での処理と同様である(ステップ1750)。また、本実施の形態におけるその他の処理も第2の実施の形態での処理と同様である。

10

【0079】

上記の処理を行うことで、第2の実施の形態での再配置判断処理312において高性能クラスに再配置先の未使用物理領域が見つからなかった場合に、高性能クラスから低性能クラスへ論理領域の再配置を、高性能クラスへの再配置に先立って行い、再配置先の未使用物理領域を高性能クラスに用意することができる。制御部300は、上記の処理を必要に応じ繰り返し行って、十分な未使用物理領域を用意することができる。

【0080】

論理領域の再配置先を低性能クラスのパリティグループ501とするため、同一負荷に対するディスク使用時間が再配置について増大し、論理領域の再配置後のディスク装置使用率が増大する可能性があるが、ディスク使用率の小さい論理領域から再配置していくようにすることで、増大の影響を最小限に抑えることができる。

20

【0081】

< 第五の実施の形態 >

本実施の形態では、クラス600の属性の1つにアクセス種別属性を設け、アクセス種別属性を用いてシーケンシャルアクセスが顕著に行われる論理領域とランダムアクセスが顕著に行われる論理領域とを、他のパリティグループ501に自動的に物理的再配置して分離するための再配置判断について説明する。

30

【0082】

本実施の形態における計算機システムは図10に示したものである。本実施の形態では、第2の実施の形態での説明に加え、制御部300が保持する下記の情報を用いる。

【0083】

本実施の形態でのクラス属性情報402の一例を図22に示す。この例では、第2の実施の形態での例に対しアクセス種別が加えられており、クラス600のアクセス種別が、例えばシーケンシャルに設定されている場合は、クラス600がシーケンシャルアクセスに好適であると設定されていることを示す。

【0084】

本実施の形態での論理領域使用状況情報403の一例を図23に示す。この例では、第2の実施の形態での例に対し、シーケンシャルアクセス率およびランダムアクセス率が加えられている。

40

【0085】

さらに、本実施の形態において制御部300は、第2の実施の形態に加え、アクセス種別基準値情報410と論理領域属性情報411を保持する。

【0086】

アクセス種別基準値情報410の一例を図24に示す。ユーザによりまたは初期条件として、アクセス種別基準値情報410には後述のアクセス種別の判定に用いる基準値が設定されている。また、論理領域属性情報411の一例を図25に示す。アクセス種別ヒントは、各論理領域について顕著に行われると期待できるアクセス種別であり、ユーザが設定

50

する。固定については後述する。

【0087】

本実施の形態での処理は、使用状況情報取得処理311および再配置判断処理312を除いては第二の実施の形態と同様である。

【0088】

本実施の形態における使用状況情報取得処理311について図26で説明する。

【0089】

制御部300は、第2の実施の形態での使用状況情報取得処理311と同様に、論理領域についてのディスク装置使用率を算出し(ステップ1800、1810)、リード/ライト処理310での使用率内容を分析して、使用率についてシーケンシャルアクセスとランダムアクセスの比率を算出し(ステップ1820)、使用率およびアクセス種別比率を論理領域使用状況情報403に記録する(ステップ1830)。また、制御部300は、第2の実施の形態と同様にパリティグループ使用率の算出と物理領域使用状況情報404への記録を行う(ステップ1840、1850)。

10

【0090】

本実施の形態における再配置判断処理312において、再配置する論理領域の選択は第2の実施の形態と同様である(ステップ1990)。再配置判断処理312での再配置先の物理領域の選択について図27で説明する。

【0091】

制御部300は、論理領域使用情報403を参照し、再配置する論理領域についてのシーケンシャルアクセス率を取得し(ステップ1910)、アクセス種別基準値情報410に設定されている基準値と比較する(ステップ1920)。シーケンシャルアクセス率が基準値より大きい場合、制御部300は、クラス属性情報402を参照し、アクセス種別がシーケンシャルと設定されているクラス600(シーケンシャルクラス)が存在するか調べる(ステップ1950)。シーケンシャルクラスが存在する場合、制御部300は、クラス構成情報401と未使用領域情報407を参照して、シーケンシャルクラスに属する再配置元以外のパリティグループ501の未使用物理領域を取得する(ステップ1960)。さらに制御部300は、各未使用物理領域について、再配置先とした場合のパリティグループ使用率の予測値を求め(ステップ1970)、未使用物理領域の中から、再配置先とした場合にシーケンシャルクラスに設定されている上限値を超えないと予測できる未使用物理領域を、再配置先の物理領域として選択し(ステップ1980)、選択結果を第2の実施の形態同様に再配置情報408に出力する(ステップ1990)。制御部300は、使用率予測値を、第2の実施の形態と同様のパリティグループ情報409と本実施の形態における論理領域使用状況情報403および物理領域使用状況情報404から算出する。

20

30

【0092】

前記の比較において、シーケンシャルアクセス率が基準値以下である場合、制御部300は、論理領域属性情報411を参照し、論理領域についてアクセス種別ヒントがシーケンシャルと設定されているか調べる(ステップ1940)。アクセス種別ヒントにシーケンシャルと設定されていた場合、上記と同様に制御部300は、シーケンシャルクラスの有無を調べ(ステップ1950)、シーケンシャルクラスが存在する場合は、シーケンシャルクラスから再配置先の物理領域を選択する(ステップ1960~1990)。

40

【0093】

前記の比較において、シーケンシャルアクセス率が前記基準値以下であり、さらにアクセス種別ヒントがシーケンシャルでなかった場合、またはシーケンシャルクラスが存在しなかった場合、制御部300は、第2の実施の形態と同様に、シーケンシャルクラス以外のクラス600から再配置先の物理領域を選択する(ステップ2000)。

【0094】

上記の処理により、同一パリティグループ501での顕著なシーケンシャルアクセスとランダムアクセスの混在に対し、各クラス600に属性として設定されたアクセス種別と使

50

用率上限値を用いて、シーケンシャルアクセスが顕著に行われる論理領域とランダムアクセスが顕著に行われる論理領域とを、異なるパリティグループ501に自動的に再配置して分離、すなわち異なるディスク装置502に分離することができ、特にランダムアクセスに対する応答性能を改善することができる。

【0095】

また、上記の処理においては制御部300は、シーケンシャルアクセスに注目して再配置による自動的分離を行うとしたが、同様にランダムアクセスに注目して前記分離を行うことも可能である。

【0096】

上記の再配置判断処理312において、再配置する論理領域を選択した時点で、制御部300が論理領域属性情報411を参照し、論理領域に固定属性が指定されている場合は、論理領域を再配置しないとすれば、ユーザが特に再配置を行いたくないと考える論理領域がある場合、固定属性を設定することで論理領域を再配置の対象外とすることができる。上記の固定属性に関する処理は論理領域属性情報411を用いることで、前述の実施の形態にも適用できる。

【0097】

【発明の効果】

ストレージサブシステムのユーザ、または保守員が、記憶領域の物理的再配置による配置最適化を行うための作業を簡便にすることができる。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態での計算機システムの構成図である。

【図2】本発明の第1の実施の形態でのリード/ライト処理310および使用状況情報取得処理311のフローチャートである。

【図3】本発明の第1の実施の形態での論理/物理対応情報400の一例を示す図である。

【図4】本発明の第1の実施の形態での再配置判断処理312のフローチャートである。

【図5】本発明の第1の実施の形態での再配置判断対象期間情報405の一例を示す図である。

【図6】本発明の第1の実施の形態での再配置情報408の一例を示す図である。

【図7】本発明の第1の実施の形態での未使用領域情報407の一例を示す図である。

【図8】本発明の第1の実施の形態での再配置実行処理313のフローチャートである。

【図9】本発明の第1の実施の形態での再配置実行時刻情報406の一例を示す図である。

。

【図10】本発明の第2の実施の形態および第五の実施の形態の計算機システムの構成図である。

【図11】本発明の第2の実施の形態での論理/物理対応情報400の一例を示す図である。

【図12】本発明の第2の実施の形態でのクラス構成情報401の一例を示す図である。

【図13】本発明の第2の実施の形態でのクラス属性情報402の一例を示す図である。

【図14】本発明の第2の実施の形態での使用状況情報取得処理311のフローチャートである。

【図15】本発明の第2の実施の形態での論理領域使用状況情報403の一例を示す図である。

【図16】本発明の第2の実施の形態での物理領域使用状況情報404の一例を示す図である。

【図17】本発明の第2の実施の形態での再配置判断処理312のフローチャートである。

。

【図18】本発明の第2の実施の形態でのパリティグループ情報409の一例を示す図である。

【図19】本発明の第2の実施の形態での再配置実行処理313のフローチャートである

10

20

30

40

50

。【図 2 0】本発明の第 3 の実施の形態での再配置判断処理 3 1 2 のフローチャートである。

。【図 2 1】本発明の第 4 の実施の形態での再配置判断処理 3 1 2 のフローチャートである。

。【図 2 2】本発明の第 5 の実施の形態でのクラス属性情報 4 0 2 の一例を示す図である。

【図 2 3】本発明の第 5 の実施の形態での論理領域使用状況情報 4 0 3 の一例を示す図である。

【図 2 4】本発明の第 5 の実施の形態でのアクセス種別基準値情報 4 1 0 の一例を示す図である。

10

【図 2 5】本発明の第 5 の実施の形態での論理領域属性情報 4 1 1 の一例を示す図である。

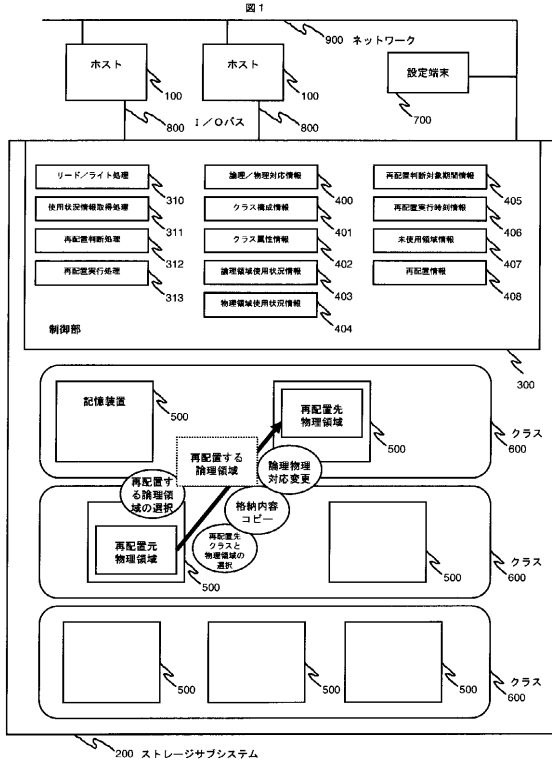
。【図 2 6】本発明の第 5 の実施の形態での使用状況情報取得処理 3 1 1 のフローチャートである。

【図 2 7】本発明の第 5 の実施の形態での再配置判断処理 3 1 2 のフローチャートである。

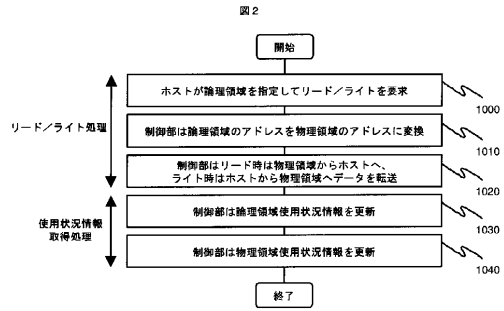
【符号の説明】

| | | |
|-------|-------------|----|
| 1 0 0 | ホスト | |
| 2 0 0 | ストレージサブシステム | |
| 2 0 1 | ディスクアレイシステム | 20 |
| 3 0 0 | 制御部 | |
| 3 1 0 | リード/ライト処理 | |
| 3 1 1 | 使用状況情報取得処理 | |
| 3 1 2 | 再配置判断処理 | |
| 3 1 3 | 再配置実行処理 | |
| 4 0 0 | 論理/物理対応情報 | |
| 4 0 1 | クラス構成情報 | |
| 4 0 2 | クラス属性情報 | |
| 4 0 3 | 論理領域使用状況情報 | |
| 4 0 4 | 物理領域使用状況情報 | 30 |
| 4 0 5 | 再配置判断対象期間情報 | |
| 4 0 6 | 再配置実行時刻情報 | |
| 4 0 7 | 未使用領域情報 | |
| 4 0 8 | 再配置情報 | |
| 4 0 9 | パリティグループ情報 | |
| 4 1 0 | アクセス種別基準値情報 | |
| 4 1 1 | 論理領域属性情報 | |
| 5 0 0 | 記憶装置 | |
| 5 0 1 | パリティグループ | |
| 5 0 2 | ディスク装置 | 40 |
| 6 0 0 | クラス | |
| 7 0 0 | 制御端末 | |
| 8 0 0 | I/Oバス | |
| 9 0 0 | ネットワーク | |

【 図 1 】



【 図 2 】

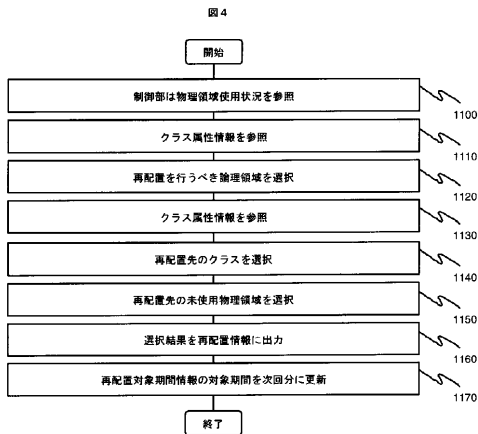


【 図 3 】

Figure 3 is a table mapping Logical Address (論理アドレス) to Physical Address (物理アドレス).

| 論理アドレス | 物理アドレス | |
|-----------|--------|-----------|
| | 記憶装置番号 | 記憶装置内アドレス |
| 0~999 | 0 | 0~999 |
| 1000~1999 | 0 | 1000~1999 |
| 2000~2999 | 1 | 0~999 |
| 3000~3999 | 1 | 1000~1999 |

【 図 4 】



【 図 6 】

Figure 6 is a table showing reconfiguration information for two examples.

| 番号 | 論理領域 | 再配置元物理領域 | | 再配置先物理領域 | |
|----|-----------|----------|-----------|----------|-----------|
| | | 記憶装置番号 | 記憶装置内アドレス | 記憶装置番号 | 記憶装置内アドレス |
| 1 | 0~999 | 0 | 0~999 | 10 | 0~999 |
| 2 | 1000~1999 | 0 | 1000~1999 | 10 | 1000~1999 |

【 図 5 】

Figure 5 is a table showing reconfiguration schedule information.

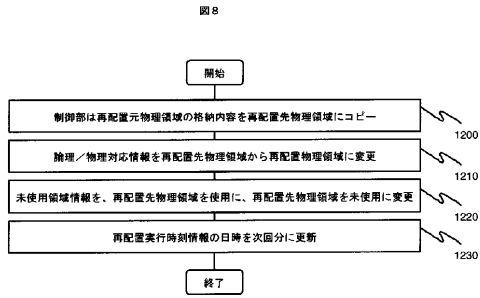
| 項目 | 内容 |
|--------|-------------------|
| 開始日時 | 1999年8月11日 8時30分 |
| 終了日時 | 1999年8月11日 17時15分 |
| 期間更新情報 | 毎日 (+24時間) |

【 図 7 】

Figure 7 is a table showing memory device status.

| 記憶装置番号 | 記憶装置内アドレス | 使用/未使用 |
|--------|-----------|--------|
| 0 | 0~999 | 使用 |
| 0 | 1000~1999 | 使用 |
| 0 | 2000~2999 | 未使用 |
| 0 | 3000~3999 | 未使用 |

【 図 8 】

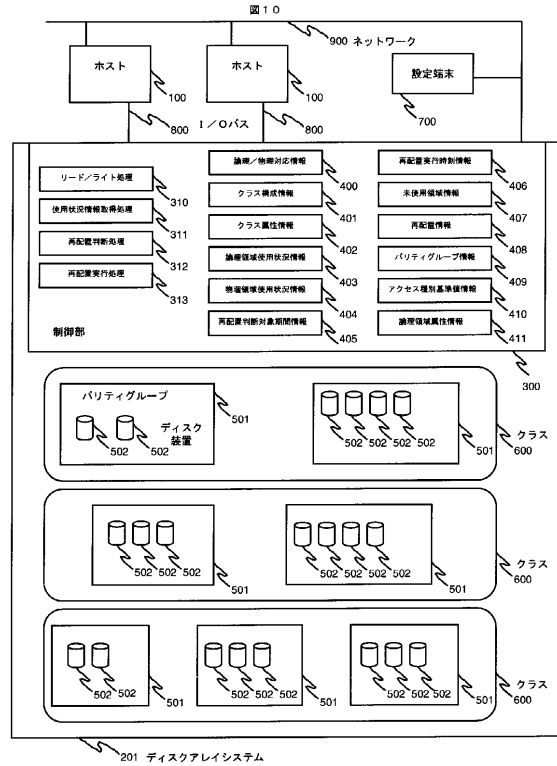


【 図 9 】

図 9

| | |
|--------|------------------|
| 日時 | 1999年8月11日 22時0分 |
| 日時更新情報 | 毎日 (+2.4時間) |

【 図 10 】

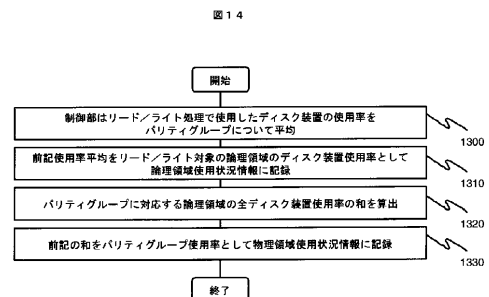


【 図 11 】

図 11

| 論理アドレス | 物理アドレス | | | | |
|-----------|------------|--------|-----------|--------|-----------|
| | パリティグループ番号 | データ | | 冗長データ | |
| | | 記憶装置番号 | 記憶装置内アドレス | 記憶装置番号 | 記憶装置内アドレス |
| 0~999 | 100 | 0 | 0~999 | 20 | 0~999 |
| 1000~1999 | 100 | 0 | 1000~1999 | 20 | 1000~1999 |
| 2000~2999 | 101 | 1 | 0~999 | 41 | 0~999 |
| 3000~3999 | 101 | 1 | 1000~1999 | 41 | 1000~1999 |

【 図 14 】



【 図 12 】

図 12

| クラス番号 | パリティグループ数 | パリティグループ番号 |
|-------|-----------|--------------------|
| 0 | 3 | 100, 110, 120 |
| 1 | 2 | 101, 111 |
| 2 | 4 | 102, 112, 122, 132 |

【 図 15 】

図 15

| 日時 | 論理アドレス | ディスク装置使用率 (%) |
|------------------|-----------|---------------|
| 1999年8月11日 8時0分 | 0~999 | 18 |
| | 1000~1999 | 32 |
| 1999年8月11日 8時15分 | 0~999 | 20 |
| | 1000~1999 | 30 |
| 1999年8月11日 8時30分 | 0~999 | 22 |
| | 1000~1999 | 28 |

【 図 13 】

図 13

| クラス番号 | 使用率上限値 (%) | クラス間性能順位 | 再配置実行上限値 (%) | 固定 |
|-------|------------|----------|--------------|----|
| 0 | 60 | 1 | 70 | - |
| 1 | 70 | 2 | 80 | 固定 |
| 2 | 80 | 3 | 90 | - |

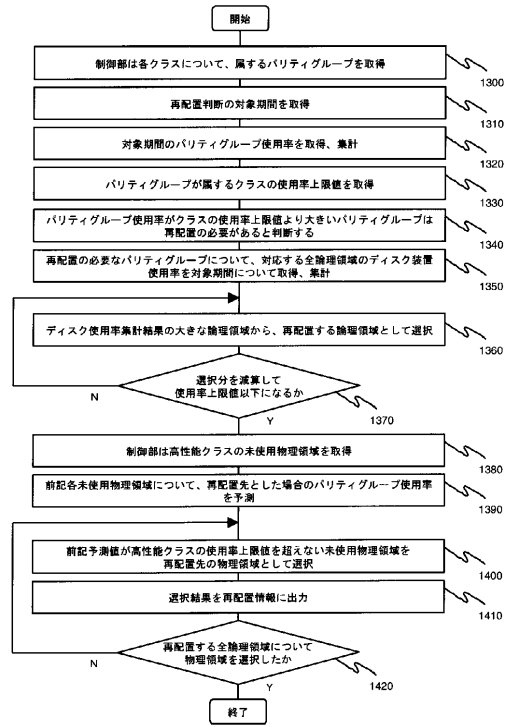
【 図 16 】

図 16

| 日時 | パリティグループ番号 | 使用率 (%) |
|------------------|------------|---------|
| 1999年8月11日 8時0分 | 100 | 68 |
| | 101 | 52 |
| 1999年8月11日 8時15分 | 100 | 70 |
| | 101 | 50 |
| 1999年8月11日 8時30分 | 100 | 72 |
| | 101 | 48 |

【 図 17 】

図 17



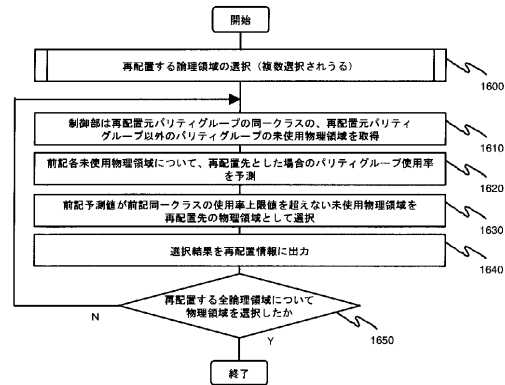
【 図 18 】

図 18

| パリティグループ番号 | RAID構成 | ディスク装置性能 | 固定 |
|------------|------------|----------|----|
| 100 | RAID5 3D1P | 110 | — |
| 101 | RAID1 1D1P | 100 | 固定 |
| 102 | RAID5 6D1P | 95 | — |

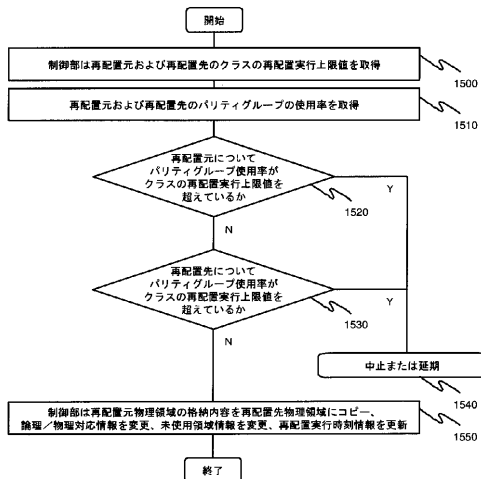
【 図 20 】

図 20

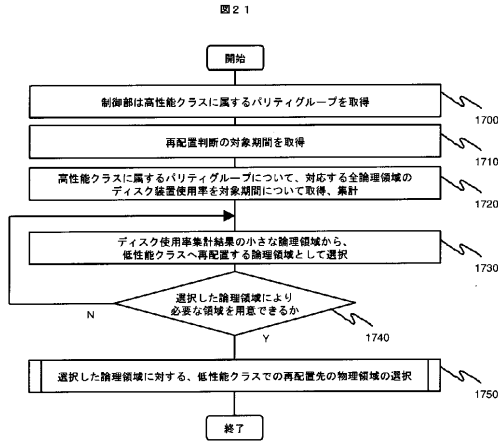


【 図 19 】

図 19



【 図 2 1 】



【 図 2 2 】

図 2 2

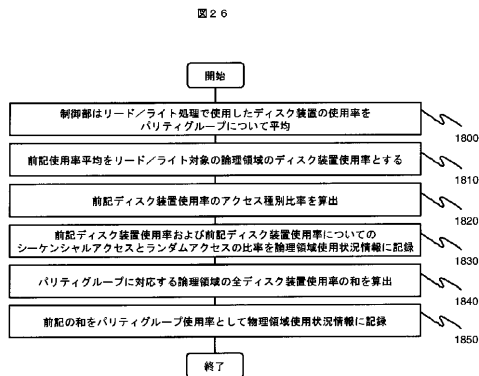
| クラス番号 | 使用率上限値 (%) | クラス間性能階位 | 再配置実行上限値 (%) | 固定 | アクセス種別 |
|-------|------------|----------|--------------|----|---------|
| 0 | 60 | 1 | 70 | — | — |
| 1 | 70 | 2 | 80 | — | — |
| 2 | 80 | 3 | 90 | — | シーケンシャル |

【 図 2 5 】

図 2 5

| 論理アドレス | アクセス種別ヒント | 固定 |
|-----------|-----------|----|
| 0~999 | — | — |
| 1000~1999 | — | — |
| 2000~2999 | シーケンシャル | — |
| 3000~3999 | — | 固定 |

【 図 2 6 】



【 図 2 3 】

図 2 3

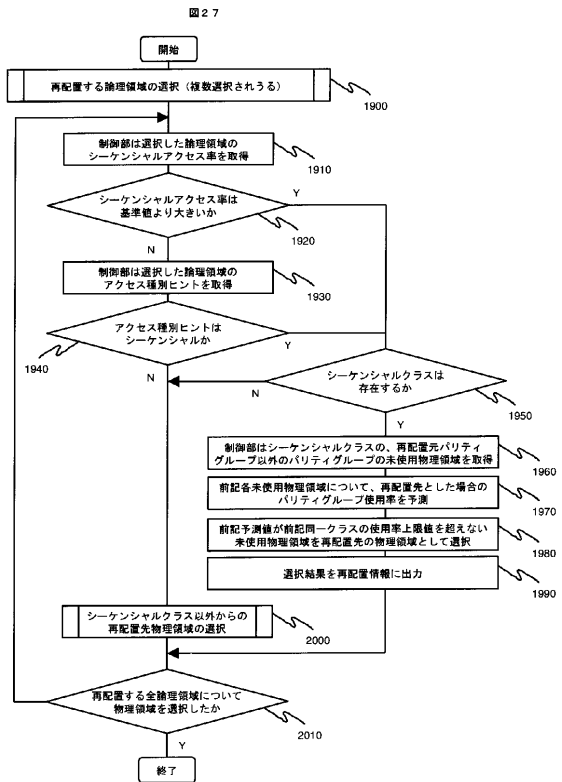
| 日時 | 論理アドレス | ディスク装置使用率 (%) | シーケンシャルアクセス率 (%) | ランダムアクセス率 (%) |
|------------------|-----------|---------------|------------------|---------------|
| 1999年8月11日 8時0分 | 0~999 | 18 | 78 | 22 |
| | 1000~1999 | 32 | 52 | 48 |
| 1999年8月11日 8時15分 | 0~999 | 20 | 80 | 20 |
| | 1000~1999 | 30 | 50 | 50 |
| 1999年8月11日 8時30分 | 0~999 | 22 | 82 | 18 |
| | 1000~1999 | 28 | 48 | 52 |

【 図 2 4 】

図 2 4

| アクセス判定基準値 (%) |
|---------------|
| 75 |

【 図 2 7 】



フロントページの続き

(72)発明者 山神 憲司

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 荒井 弘治

神奈川県小田原市国府津2880番地 株式会社日立製作所 ストレージシステム事業部内

審査官 藤井 浩

(56)参考文献 特開平04-151745(JP,A)

特開平11-085411(JP,A)

特開平07-056691(JP,A)

特開平09-274544(JP,A)

特開平11-095934(JP,A)

(58)調査した分野(Int.Cl.⁷, DB名)

G06F 3/06 - 3/08

G06F 12/00