(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3544835号
(P3544835)
(45)発行日　平成16年7月21日(2004.7.21)　(24)登録日　平成16年4月16日(2004.4.16)

(51)Int.Cl.[7]　F I
G06F 3/06　G06F 3/06　540
G06F 12/00　G06F 3/06　302J
G06F 12/00　501B

請求項の数 6　(全 14 頁)

(21)出願番号　特願平9-256315
(22)出願日　平成9年9月22日(1997.9.22)
(65)公開番号　特開平11-95934
(43)公開日　平成11年4月9日(1999.4.9)
審査請求日　平成13年4月26日(2001.4.26)

前置審査

(73)特許権者　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地
(74)代理人　100080001
弁理士　筒井　大和
(72)発明者　黒川　勇
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
(72)発明者　竹内　久治
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

最終頁に続く

(54)【発明の名称】記憶システムの制御方法および制御装置

(57)【特許請求の範囲】
【請求項１】
複数の物理デバイスを含む複数のデバイスグループの各々に外部処理装置からの入出力要求の対象となる論理デバイスが定義される記憶システムにおける制御方法であって、
互いに異なるデバイスグループに属し、前記論理デバイスに対応する前記複数の物理デバイス毎に複数の論理範囲を設け、前記外部処理装置からの前記入出力要求の頻度が異なる論理範囲を交換するものであり、
前記複数のデバイスグループのうち前記入出力要求回数の総和が最大のデバイスグループに属し入出力要求回数が最大の論理範囲と、前記複数のデバイスグループのうち前記入出力要求回数の総和が最小のデバイスグループに属し入出力要求回数が最小の論理範囲とを交換する場合、
作業領域を設定するステップと、前記入出力要求回数の総和が前記最大のデバイスグループに属し前記入出力要求回数が前記最大の論理範囲または前記入出力要求回数の総和が前記最小のデバイスグループに属し前記入出力要求回数が前記最小の論理範囲のうち、一方のデバイスグループに属した一方の論理範囲を前記作業領域に移動させるステップと、他方のデバイスグループに属した他方の論理範囲を前記一方のデバイスグループ内に移動させるステップと、前記作業領域の前記一方の論理範囲を前記他方のデバイスグループ内に移動させるステップとを実行し、
移動中の前記一方または前記他方の論理範囲に対して前記外部処理装置から入力（データ読み出し）要求が発生した場合には、移動前の前記一方または前記他方の論理範囲または

前記作業領域からデータを読み出して応答し、
移動中の前記一方または前記他方の論理範囲に対して前記外部処理装置から出力（データ書き込み）要求が発生した場合には、移動前の前記一方または前記他方の論理範囲または前記作業領域と、移動先の前記作業領域または前記一方または前記他方のデバイスグループの両方にデータを書き込むことを特徴とする記憶システムの制御方法。
【請求項２】
請求項１記載の記憶システムの制御方法において、
複数の前記デバイスグループおよび前記論理範囲における前記入出力要求の受付頻度の観測結果を可視化することを特徴とする記憶システムの制御方法。
【請求項３】
請求項１記載の記憶システムの制御方法において、
前記作業領域を、キャッシュメモリに有することを特徴とする記憶システムの制御方法。
【請求項４】
複数の物理デバイスを含む複数のデバイスグループの各々に外部処理装置からの入出力要求の対象となる論理デバイスが定義される記憶システムにおける制御装置であって、
互いに異なるデバイスグループに属し、前記論理デバイスに対応する前記複数の物理デバイス毎に複数の論理範囲を設け、前記外部処理装置からの前記入出力要求の頻度が異なる論理範囲を交換する手段を有するものであり、
前記手段は、
前記複数のデバイスグループのうち前記入出力要求回数の総和が最大のデバイスグループに属し入出力要求回数が最大の論理範囲と、前記複数のデバイスグループのうち前記入出力要求回数の総和が最小のデバイスグループに属し入出力要求回数が最小の論理範囲とを交換する場合、
作業領域を設定するステップと、前記入出力要求回数の総和が前記最大のデバイスグループに属し前記入出力要求回数が前記最大の論理範囲または前記入出力要求回数の総和が前記最小のデバイスグループに属し前記入出力要求回数が前記最小の論理範囲のうち、一方のデバイスグループに属した一方の論理範囲を前記作業領域に移動させるステップと、他方のデバイスグループに属した他方の論理範囲を前記一方のデバイスグループ内に移動させるステップと、前記作業領域の前記一方の論理範囲を前記他方のデバイスグループ内に移動させるステップとを実行し、
移動中の前記一方または前記他方の論理範囲に対して前記外部処理装置から入力（データ読み出し）要求が発生した場合には、移動前の前記一方または前記他方の論理範囲または前記作業領域からデータを読み出して応答し、
移動中の前記一方または前記他方の論理範囲に対して前記外部処理装置から出力（データ書き込み）要求が発生した場合には、移動前の前記一方または前記他方の論理範囲または前記作業領域と、移動先の前記作業領域または前記一方または前記他方のデバイスグループの両方にデータを書き込むことを特徴とする記憶システムの制御装置。
【請求項５】
請求項４記載の記憶システムの制御装置において、
前記手段は、
複数の前記デバイスグループおよび前記論理範囲における前記入出力要求の受付頻度の観測結果を可視化することを特徴とする記憶システムの制御装置。
【請求項６】
請求項４記載の記憶システムの制御装置において、
前記作業領域を、キャッシュメモリに有することを特徴とする記憶システムの制御装置。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】
本発明は、外部記憶システムの論理デバイスデータ交換技術に関し、特に、ディスクアレイ等のアレイ型外部記憶装置の制御等に適用して有効な技術に関する。

【０００２】
【従来の技術】
一般に、アレイ型外部記憶装置は、複数の物理デバイスの集まりを１アレイグループとし、アレイグループ内では、物理デバイスに対する入出力要求の負荷分散を目的に、論理デバイスを分割し、分割したそれぞれの領域を各物理デバイスに分散させている。この為、各物理デバイスを並列に処理させることができ、大容量の１ファイル一つのみが、１論理デバイスに格納されている場合では、大容量のファイルの入出力を高速に処理することができる。
【０００３】
しかし、複数の小容量のファイルが、ある１論理デバイスに集中して格納されている情報処理システムにおいて、前述の複数の小容量のファイルのそれぞれに同時に入出力要求が生じた場合、たとえアレイ型外部記憶装置を用いていても、上位処理装置の各論理デバイスに対応した入出力制御部がビジーとなり、システム全体の処理能力が低下するという技術的課題が発生する。
【０００４】
この技術的課題を解決する方法として特開平０７－１５２４９８号公報（以下、第１の従来技術と記す）に開示された技術が挙げられる。すなわち、この第１の従来技術は、一つのファイルが一つの論理デバイスに格納されるようにするというものであり、本技術を適用すれば、上位処理装置の各論理デバイスに対応した入出力制御部の負荷が分散され、前述の技術的課題となった入出力制御部のビジーは解消される。さらに、アレイグループ内では、各論理デバイスの容量を小さく定義すれば、論理デバイス数をほぼ無制限に定義可能であり、この第１の従来技術をアレイ型外部記憶装置に適用すると、ファイルの容量に一致させる容量の論理デバイスを定義することになり、少数のアレイグループで対応できる。
【０００５】
しかしながら、アレイグループ内に多数の論理デバイスすなわちファイルが格納された情報処理システムにおいて、小容量のファイルのそれぞれに同時に入出力要求が生じた場合、アレイグループを形成する各物理デバイスに対する入出力要求の頻度は高くなる。最悪、同時に複数の入出力要求が一つの物理デバイスに集中し、逆に今度は物理デバイスのビジーが発生する。やはりこれもシステム全体の処理能力が低下する問題となる。
【０００６】
この問題を決する方法として、さらに前述の第１の従来技術では、アレイグループ内の論理デバイスすなわちファイルの入出力要求の頻度を考慮して、ファイルをアレイグループ内の別の物理デバイスに格納させて、アレイグループ内の各物理デバイスの負荷を均等化させようとする技術が開示されている。
【０００７】
この第１の従来技術に類似の公知例として、特開平０８－２０２５０３号公報（以下、第２の従来技術と記す）に開示された技術が挙げられる。この第２の従来技術では、ディスクアレイ装置において、同時にアクセスされる複数のデータファイルの組み合わせ回数を記録し、当該組み合わせ回数の多い複数のデータファイルを異なる物理デバイスに分散させることで、特定の物理デバイスにアクセスが集中することを回避しようとするものである。
【０００８】
【発明が解決しようとする課題】
物理デバイスの負荷分散を図る為の、前述の第１の従来技術は負荷分散を可能とし非常に有効である。しかしながら、実施において、アレイグループ内の論理デバイスすべてに対して均等に入出力要求の頻度が高い場合では、各物理デバイスに対する入出力要求の頻度も均等に高いものとなり、アレイグループ内で論理デバイスの分散を行なっても、何処も均等に高い入出力要求の頻度なので、当該従来技術を適用しても各物理デバイスに対する入出力要求の頻度は変化しないという現象が発生し、物理デバイスのビジーが発生しうる

。すなわち、物理デバイスのビジーが回避できず、システムの処理能力は上がらない。従って、何等かの別の方法で、この物理デバイスのビジーを解消する必要がある。

【０００９】

本発明の目的は、物理デバイス上に複数の論理デバイスを設定して稼動する情報処理システムにおける論理デバイスの入出力性能の向上を実現することにある。

【００１０】

本発明の他の目的は、物理デバイス上に複数の論理デバイスを設定して稼動する情報処理システムにおける物理デバイスの可用性の向上を実現することにある。

【００１１】

本発明の他の目的は、物理デバイス上に複数の論理デバイスを設定して稼動する情報処理システムにおいて、利用者の要請に合わせた論理デバイスの設定の効率化および多様化を実現することにある。

【００１２】

【課題を解決するための手段】

本発明は、各々が少なくとも一つの物理デバイスを含む複数のデバイスグループの各々に、任意の外部処理装置から見て入出力要求の対象となる少なくとも一つの論理デバイスが定義されてなる情報処理システムの制御方法であって、入出力要求の頻度が高いデバイスグループ内の入出力要求の頻度が高い物理デバイスに存在する論理デバイス、もしくは、論理デバイスの一部を、他のデバイスグループに移動するか、または他のデバイスグループの論理デバイスと交換し、デバイスグループ内の物理デバイスの使用負荷を低減させかつデバイスグループ間の負荷を均等とすることにより、外部処理装置からの入出力要求を高速に処理可能にするものである。

【００１３】

この場合、論理デバイスの移動や入れ換えの契機は、観測される入出力要求の受付頻度等の統計情報に基づいて、情報処理システムが自律的かつ自動的に実行してもよいし、外部のサービス機器等からの指示にて任意の契機で任意の論理デバイスを移動または入れ替えるようにしてもよい。

【００１４】

より具体的には、一例として、デバイスグループを構成する物理デバイス群と、これらの物理デバイス群と外部処理装置との間に介在し、物理デバイス群を複数の論理デバイスとして機能させる制御動作を行う制御装置とを備えた情報処理システムにおいて、以下のような手段を設ける。

【００１５】

すなわち、自動的に二つ論理デバイスの交換を実現する手段として、制御装置は、外部処理装置から制御装置の配下の各論理デバイスに対する入出力要求の際に、入出力要求対象の論理デバイスが在るデバイスグループ内の各物理デバイスに対する入出力要求の頻度を表わす情報として入出力回数を論理デバイス単位に常時インクリメントし制御情報用メモリに記憶していき、ある周期で、前述の入出力回数を全論理デバイス分参照し、各論理デバイスが存在するデバイスグループに対する入出力要求の頻度を表わす情報としてデバイスグループ内での前述の入出力回数の総和値と制御装置の配下の総和値の平均値を求め、前述の総和値の最大値が平均値より極端に大きいか、もしくは、前述の総和値の最小値が平均値より極端に小さい場合に、前述の総和値が最大のデバイスグループ内において前述の入出力回数が最大の論理デバイスと、前述の総和値が最小のデバイスグループにおいて前述の入出力回数が最小の論理デバイスとの交換を行う手段を提供する。

【００１６】

また、制御装置は、二つの論理デバイスの交換を実現する手段として、第一に、制御装置の配下の空き領域が存在するデバイスグループを作業領域とし、任意の論理デバイスをこの制御装置の配下の空き領域が存在するデバイスグループへ移動し定義する手段を提供する。第二に、交換する二つの論理デバイスのどちらか一方を、前述の第一の手段を用いて、空き領域が存在するデバイスグループへの移動を完了させ、一方の論理デバイス移動後

の空き領域を作業領域とし、同様に他方の論理デバイスを前述の第一の手段を用いて、一方の論理デバイス移動後の空き領域への移動を完了させ、更に他方の論理デバイス移動後の空き領域を作業領域とし、先に移動した論理デバイスを更に前述の第一の手段を用いて、他方の論理デバイス移動後の空き領域への移動を完了させることで交換を行う手段を提供する。

【００１７】

第一の手段は、デバイスグループ内の論理デバイスを分割した各分割領域の位置を示す配置情報テーブルを該当デバイスグループ内の各論理デバイスに対応する形態で記憶し、この配置情報テーブルを指し示す情報であるデバイスグループ間の変換テーブルを各論理デバイスに対応する形態で、かつ、各論理デバイスが現存している該当デバイスグループ内の配置情報テーブルのみを指し示す情報で記憶しておく。

【００１８】

そして、任意の論理デバイスを制御装置の配下の空き領域が存在するデバイスグループへ移動する際に、移動対象論理デバイスの移動先デバイスグループ内での新しい配置情報テーブルを記憶しておき、前述の変換テーブルの情報を、現存している該当デバイスグループ内の配置情報テーブルと移動先デバイスグループ内の新しい配置情報テーブルの両方を指し示す情報に更新して記憶しておき、前述の新旧の配置情報テーブルの情報を参照し、移動対象論理デバイスのデータの移動先を決定し、移動対象論理デバイスのデータを、直接にまたはキャッシュメモリを介し移動先デバイスグループ内の各物理デバイスに転送し、移動対象論理デバイスのデータの移動が完全に終了した際に、前述の変換テーブルの情報を、移動先デバイスグループ内の新しい配置情報テーブルしか参照できない内容の情報に更新して記憶し、移動完了とする。

【００１９】

また、移動対象論理デバイスの移動完了前に、外部処理装置からの移動対象論理デバイスに対する入力要求では、変換テーブルの情報から現存しているデバイスグループ内の配置情報テーブルを参照して、要求に該当するデータを外部処理装置へ転送する。また、外部処理装置からの移動対象論理デバイスに対する出力要求では、変換テーブルの情報から現存している該当デバイスグループ内の配置情報テーブルを参照して、外部処理装置から転送されるデータを、移動対象論理デバイスへ転送し、かつ、変換テーブルの情報から移動先デバイスグループ内の配置情報テーブルを参照して、移動先デバイスグループ内の各物理デバイスに転送するという動作を行うものである。

【００２０】

なお、制御装置は、論理デバイスをある大きさの論理範囲で区分し、この論理範囲を論理デバイスの代替とし、上述と同様に異なるデバイスグループ間における移動や交換を行うこともできる。

【００２１】

【発明の実施の形態】

以下、本発明の実施の形態を図面を参照しながら詳細に説明する。

【００２２】

図１は、本発明の一実施の形態である論理デバイスデータ交換方法を実施する外部記憶システムを含む情報処理システムの構成の一例を示す概念図であり、図２は、その構成の一部を取り出して示す概念図、図３は、本実施の形態の外部記憶システムの論理デバイスデータ交換方法にて用いられる制御情報の一例を示す概念図である。また、図４、図５および図６は、本実施の形態の外部記憶システムの論理デバイスデータ交換方法の作用の一例を示すフローチャート、図７は、本実施の形態の情報処理システムの仕様の一例を示す説明図である。

【００２３】

（Ａ）本実施の形態の情報処理システムの仕様例

図７に、本発明の論理デバイス、もしくは、論理デバイス内のある一部の移動・交換の一実施の形態の仕様一覧表を示す。

【００２４】
（Ｂ）本実施の形態の情報処理システムの全体構成の説明
図１に示す本実施の形態の情報処理システムは、情報の演算等を行う処理装置１０と、データ転送等の入出力制御を行うチャネル１１と、アレイ型外部記憶装置１４と、これら装置の制御を行う制御装置１２とから構成されている。
【００２５】
また制御装置１２は、制御を実施する際に使用する制御情報を記憶する制御情報用メモリ１２２に対する入出力機能と、制御情報用メモリ１２２に格納されている入出力情報をたとえば図８に例示されるように視覚的に表示できる機能と、制御装置１２に対して論理デバイスの交換命令機能を持つサービス機器１３を備え、かつ内部時計を持ち時刻を取得できる。このシステムは処理装置１０の命令によりチャネル１１、制御装置１２のキャッシュメモリ１２３を介してアレイ型外部記憶装置１４内に設定されている論理デバイス１４３に対してデータの記録再生を行うものである。尚、システムにおいて、１論理デバイスには固有の論理デバイス識別子（以下、論理デバイス番号３４と称す）が付与されており、また、制御装置１２の配下において、一つの物理デバイスには固有の物理デバイス識別子（以下、物理デバイス番号と称す）および一つのアレイグループ（デバイスグループ）には固有のアレイグループ識別子（以下、アレイグループ番号３６と称す）が定義される。
【００２６】
一方、制御装置１２の配下のアレイ型外部記憶装置１４は図２に例示されるように、複数の物理デバイス１４２の集まりを一つのアレイグループとし、複数のアレイグループ１４１より構成されている。個々のアレイグループ１４１内には処理装置１０のデータの入出力要求対象となる複数の論理デバイス１４３が、制御装置１２により、１論理デバイスを制御上の論理範囲で分割した形態で定義され、それぞれの分割領域２０が各物理デバイス１４２に分散している。処理装置１０の命令により論理デバイス１４３に対してのデータの記録再生が発生した場合、制御装置１２の制御部１２１は、論理デバイス１４３の記録再生対象データが存在する分割領域２０の在る物理デバイス１４２を対象に入出力を行う。
【００２７】
（Ｃ）制御装置配下のアレイグループ内の論理デバイス定義に関する説明
本実施の形態に示したサービス機器１３より、一つのアレイグループ１４１内に設定する論理デバイス数３７と論理デバイス番号３４と論理デバイス容量３５と定義対象のアレイグループ番号３６が、制御部１２１が使用する制御情報用メモリ１２２上の論理デバイス管理テーブル３３に入力される。制御装置１２の制御部１２１は、これら論理デバイス１４３の各々に対して、１論理デバイスを制御上の論理範囲で分割し、分割した領域をアレイグループ１４１内の各物理デバイス１４２に分散定義し、論理デバイス１４３の各分割領域２０の位置を示す配置情報テーブル４０をアレイグループ単位に各論理デバイス１４３に対応する形態で制御情報用メモリ１２２上に記憶し、この配置情報テーブル４０を指し示す情報（以下、配置情報テーブルポインタと称す）を論理デバイス管理テーブル３３内のアレイグループ間変換テーブル３８に記憶する。この時の配置情報テーブルポインタは各論理デバイス１４３が現存している該当アレイグループの配置情報テーブル４０を指し示している。尚、制御装置１２は装置立ち上げ時に接続しているアレイグループ１４１の、全アレイグループ番号を接続アレイグループ番号テーブル４１、物理デバイス容量４２を、制御情報用メモリ１２２に記憶している。制御情報用メモリ１２２に設定される、論理デバイス管理テーブル３３、配置情報テーブル４０、アレイグループ間変換テーブル３８、接続アレイグループ番号テーブル４１、の構成の一例を図３に示す。
【００２８】
（Ｄ）論理デバイスの他のアレイグループへの移動処理に関する説明
本実施の形態に示した制御装置１２の制御部１２１は、当該制御装置１２の配下に、空き領域が存在するアレイグループ１４１が在るかを判断する。すなわち、制御装置１２の配

下に定義されているアレイグループ番号３６のうち全論理デバイス管理テーブル３３に設定されていない番号があれば、当該番号のアレイグループ１４１は未使用ということであり、また、アレイグループ１４１を構成する物理デバイス１４２の容量の総和からアレイグループ１４１内の論理デバイス容量３５の総和を引いた差が、移動対象論理デバイスの論理デバイス容量３５より等しいかもしくは大きければ、移動可能な空き領域が存在することになる。移動可能であるならば、制御装置１２の制御部１２１は、移動先のアレイグループ１４１内での配置情報テーブル（以下、新配置情報テーブルと称す）を定義して記憶しておき、アレイグループ間変換テーブル３８に、新たに新配置情報テーブルを指し示す配置情報テーブルポインタ（以下、新配置情報テーブルポインタ３８ａと称す）を記憶する。この時点で、アレイグループ間変換テーブル３８には新配置情報テーブルポインタ３８ａと移動元の配置情報テーブルポインタ（以下、旧配置情報テーブルポインタ３８ｂと称す）が存在する。制御装置１２の制御部１２１は、旧配置情報テーブルの情報が示す分割領域２０のデータを対応する物理デバイス１４２より、直接にまたはキャッシュメモリ１２３を介し新配置情報テーブルの情報が示す分割領域２０が在る物理デバイス１４２に転送する。移動対象の論理デバイス１４３の全分割領域の移動が完全に終了した際に、アレイグループ間変換テーブル３８に記憶した旧配置情報テーブルポインタ３８ｂを、新配置情報テーブルポインタ３８ａで更新して記憶し、新配置情報テーブルポインタ３８ａは消去し、移動完了とする。

【００２９】

また、移動対象の論理デバイス１４３の移動完了前に、処理装置１０からの当該移動対象の論理デバイス１４３に対する入力（データ読み出し）要求では、アレイグループ間変換テーブル３８の旧配置情報テーブルポインタ３８ｂの指す配置情報テーブル４０を参照して、要求に該当するデータが存在する物理デバイス１４２よりデータを読み出して処理装置１０へ転送する。

【００３０】

同様に、処理装置１０からの移動対象の論理デバイス１４３に対する出力（データ書き込み）要求では、アレイグループ間変換テーブル３８の旧配置情報テーブルポインタ３８ｂの指す配置情報テーブル４０を参照して、処理装置１０から転送されるデータを、当該出力要求に該当するデータが存在する物理デバイス１４２へ転送して書き込み、かつ、アレイグループ間変換テーブル３８の新配置情報テーブルポインタ３８ａの指す配置情報テーブル４０を参照して、移動先アレイグループ内の物理デバイス１４２にも転送して書き込む。

【００３１】

この動作をフローチャートで示したのが、図４、図５である。図４のフローチャートでは、制御部１２１が一つの論理デバイス１４３の移動を実施するにあたり、ステップ４０１で、移動先の存在の確認を行う。すなわち、論理デバイス管理テーブル３３内のアレイグループ番号３６に記憶されておらず接続アレイグループ番号テーブル４１に記憶されているアレイグループ１４１を移動先のアレイグループ１４１とする。これがステップ４０２である。次にステップ４０３で、移動先のアレイグループ１４１での移動対象の論理デバイス１４３の新配置情報テーブルを作成する。ステップ４０４で、新配置情報テーブルを参照できるよう、アレイグループ間変換テーブル３８に新配置情報テーブルポインタ３８ａを記憶する。ステップ４０５では、旧配置情報テーブルが示す分割領域のデータをそれぞれ、新配置情報テーブルが示す分割領域へ転送する。ステップ４０６で全分割領域の移動が完了したことを確認した後、ステップ４０７でアレイグループ間変換テーブル３８の旧配置情報テーブルポインタ３８ｂを新配置情報テーブルポインタ３８ａで更新し記憶する。以上で論理デバイス１４３の移動が完了する。

【００３２】

図５のフローチャートでは、ステップ５０１で処理装置１０より入出力要求がされたことをチェックし、入出力要求がされたならば、ステップ５０２において、処理装置１０の入出力要求対象の論理デバイス１４３のアレイグループ間変換テーブル３８に値が記憶され

ていることをチェックする。値があった場合、ステップ５０３で、旧配置情報テーブルポインタ３８ｂの指す旧配置情報テーブルを参照し、入出力要求対象の分割領域がどこに在るかを識別する。ステップ５０４では処理装置１０が入力要求を行なっているか否かを判定する。入力要求（データ読み出し）の場合は、ステップ５０５に進み、識別した処理対象の分割領域の在る物理デバイス１４２より処理装置１０にデータを転送する。他方、ステップ５０４で入力要求でない場合（出力要求（データ書き込み）の場合）は、ステップ５０６に進み、識別した処理対象の分割領域の在る物理デバイス１４２へ処理装置１０からの書き込みデータを転送する。次にステップ５０７で、新配置情報テーブルポインタ３８ａより新配置情報テーブルを参照し、出力要求対象の分割領域がどこに在るかを識別する。その後、ステップ５０８で、識別した処理対象の分割領域が定義されている移動先アレイグループ内の物理デバイス１４２へ処理装置１０からのデータを転送する。

【００３３】

（Ｅ）論理デバイスと他のアレイグループ内の論理デバイスの交換処理に関する説明

本実施の形態に示した制御装置１２の制御部１２１は、交換する二つの論理デバイス１４３のどちらか一方の論理デバイスを、前記（Ｄ）の操作を用いて、空き領域が存在するアレイグループ１４１への移動を完了させ、さらに、他方の論理デバイスを前記（Ｄ）の操作の変形の形で、すなわち、空き領域が存在するアレイグループ１４１への移動とするのではなく、代替として、一方の論理デバイス移動後の空き領域への移動として完了させ、最後に、先に移動した論理デバイスを前記（Ｄ）の操作の変形の形で、すなわち、空き領域が存在するアレイグループ１４１への移動とするのではなく、代替として、他方の論理デバイス１４３の移動後の空き領域への移動を完了させることで交換を行うというものである。

【００３４】

（Ｆ）制御装置による二つ論理デバイスの自動交換処理に関する説明

本実施の形態に示した制御装置１２の制御部１２１は、制御装置１２の制御部１２１立ち上げ後より、処理装置１０から制御装置１２の配下の論理デバイス１４３に対し入出力要求の度に、入出力要求対象の論理デバイス１４３が在るアレイグループ１４１内の各物理デバイス１４２に対する入出力回数を、論理デバイス管理テーブル３３内に設けられた各物理デバイス毎および各論理デバイス毎の入出力要求回数３９のエントリをインクリメントし記憶する。

【００３５】

まず、制御装置１２の制御部１２１立ち上げ後、最初の入出力要求があった時刻を制御情報用メモリ１２２上の開始時刻テーブル３１に記憶し、以後、処理装置１０から制御装置１２の配下の論理デバイス１４３に対し入出力要求の度に、現在の時刻と開始時刻テーブル３１に記憶した時刻を比較し、その差が、ある一定時間を超えた場合に、記憶している入出力要求回数３９の総和を各論理デバイス１４３が存在するアレイグループ１４１毎の入出力要求の回数とし、制御装置１２の配下に存在する全アレイグループ１４１中、アレイグループ１４１毎の入出力要求回数３９の総和が最大のアレイグループ１４１内の入出力要求回数３９が最大の論理デバイス１４３と全アレイグループ１４１中、アレイグループ１４１毎の入出力要求回数３９の総和が最小のアレイグループ１４１内の入出力要求回数３９が最小の論理デバイス１４３の交換を実施すると決定する。

【００３６】

一方、サービス機器１３により論理デバイス自動交換機能が有効と入力されているならば、制御情報用メモリ１２２上の自動交換機能有効フラグ３２がＯＮとなっている。すなわち、交換を実施すると決定した際に、自動交換機能有効フラグ３２がＯＮであれば、制御装置１２が決定した二つの論理デバイス１４３に対して、前記（Ｅ）の方法で、論理デバイス１４３の交換処理を行う。交換完了後、全論理デバイス１４３の入出力要求回数３９をクリアし、かつ、現在時刻で開始時刻テーブル３１を更新する。

【００３７】

この動作をフローチャートで示したのが図６である。この図６のフローチャートでは、制

御装置１２が立ち上がった後、ステップ６０１で開始時刻テーブル３１に現在時刻を制御装置１２内の内部時計から取得し記憶する。ステップ６０２では制御装置１２が処理装置１０からの入出力処理要求を受けたか否かを判断し、受けたのであればステップ６０３へ進み、そうでなければステップ６０２の入り口へ戻る。ステップ６０３では、入出力対象の論理デバイス１４３の在る各物理デバイス１４２に対する入出力を行ったか否かを判断し、行なっているのであればステップ６０４へ進み、そうでなければステップ６０２の入り口へ戻る。ステップ６０４では、入出力対象の論理デバイス１４３の在るアレイグループ１４１に対する入出力要求回数３９をインクリメントして記憶し、ステップ６０５では開始時刻テーブル３１の開始時刻から現在時刻の差が、所望の規定値（たとえば、１週間）より大きいか否かを判断し、大きいのであるならステップ６０６へ進み、そうでないのであればステップ６０２の入り口へ戻る。ステップ６０６では、サービス機器１３から、あらかじめ設定されている自動交換機能有効フラグ３２がＯＮであるか否かを判断し、ＯＮであるのならばステップ６０７へ進み、ＯＦＦならばステップ６０２の入り口へ戻る。ステップ６０７では、アレイグループ１４１毎の各物理デバイス１４２に対する入出力要求回数３９の制御装置１２配下の全体の平均値Ａに比べ、各物理デバイス１４２に対する入出力要求回数３９の総和の最大値Ｍが１０％以上大きいか否かを判断し、大きいのならばステップ６０９へ進み、そうでないのであればステップ６０８へ進む。ステップ６０８では、各物理デバイス１４２に対する入出力要求回数３９の総和の最小値ｍに比べ、アレイグループ１４１毎の各物理デバイス１４２に対する入出力要求回数３９の制御装置１２配下の全体の平均値Ａが、所望の規定値（たとえば、１０％）以上大きいか否かを判断し、大きいのならばステップ６０９へ進み、そうでないのであればステップ６１０へ進む。ステップ６０９では、各物理デバイス１４２に対する入出力要求回数３９の総和が最大のアレイグループ１４１内の入出力要求回数３９が最大の論理デバイス１４３と、物理デバイス１４２に対する入出力要求回数３９の総和が最小のアレイグループ１４１内の入出力要求回数３９が最小の論理デバイス１４３の交換を実施する。交換完了後、ステップ６１０ではすべての物理デバイス１４２に対する入出力要求回数３９を全てゼロで初期化する。ステップ６１１では、開始時刻テーブル３１を現在時刻で更新する。

【００３８】

（Ｇ）サービス機器による二つの論理デバイスの交換処理の起動に関する説明

本実施の形態に示した制御装置１２の制御部１２１は、サービス機器１３により、交換対象の二つの論理デバイス番号３４が制御情報用メモリ１２２上の交換対象論理デバイス番号登録テーブル３０に入力される。この操作は、たとえば図８に例示されるように、サービス機器１３の操作画面１３ａに可視化して表示された、制御情報用メモリ１２２の各種テーブル（図３）の内容を参照して、操作者が、交換対象の論理デバイス１４３を適宜選択することによって行われる。

【００３９】

すなわち、操作画面１３ａには、たとえば、図８に例示されるように、各アレイグループ１４１に対応したボックス１３ｂと、その内部に配列された当該アレイグループ１４１内の各論理デバイス１４３に対応した棒グラフ１３ｃが表示される。棒グラフ１３ｃの高さは、当該論理デバイス１４３に対する入出力要求の受付頻度を示す。ボックス１３ｂの高さは、対応するアレイグループ１４１を構成する複数の物理デバイス１４２に対する入出力要求の受付頻度の総和を示す。そして、サービス機器１３の操作者は、たとえば、最も高さの高いボックス１３ｂのアレイグループ１４１内の最も高さの高い棒グラフ１３ｃに対応した論理デバイス１４３を、最も高さの低いボックス１３ｂのアレイグループ１４１内の最も高さの低い棒グラフ１３ｃに対応した論理デバイス１４３と交換するように、交換対象論理デバイス番号登録テーブル３０に設定する。

【００４０】

この時、同時に、上述の論理デバイス自動交換処理を抑止するために、制御情報用メモリ１２２上の自動交換機能有効フラグ３２もＯＦＦされる。制御装置１２の制御部１２１は、処理装置１０から制御装置１２の配下の論理デバイス１４３に対し入出力要求の度に、

この交換対象論理デバイス番号登録テーブル３０に論理デバイス番号３４が登録されているか否かを判断している。登録されていると判断した時、登録されている二つの論理デバイス１４３が異なるアレイグループ１４１に属しているか否かをそれぞれのアレイグループ番号３６から判断し、異なるアレイグループ１４１に属している場合に、登録されている二つの論理デバイス１４３に対して、前記（Ｅ）の方法で論理デバイスの交換処理を行う。

【００４１】

（Ｈ）制御装置による二つの論理範囲の自動交換処理に関する説明

本実施の形態に示した制御装置１２では、一つの論理範囲は、一つの分割領域２０に対応しており、前記（Ｆ）においての入出力要求対象の論理デバイス１４３が在るアレイグループ１４１内の各物理デバイス１４２に対する入出力回数を、論理デバイス管理テーブル３３内に設けられた入出力要求回数３９を各物理デバイス毎にインクリメントすることで記憶する際に、物理デバイス１４２毎の入出力要求回数３９を更に小さな単位の各論理範囲毎にインクリメントし記憶しておけば、論理的に（Ｆ）と同じ操作が可能である。すなわち、記憶している入出力要求回数３９の総和を各論理デバイス１４３が存在するアレイグループ１４１毎の入出力要求の回数とし、制御装置１２の配下に存在する全アレイグループ１４１中、アレイグループ１４１毎の入出力要求回数３９の総和が最大のアレイグループ１４１内の入出力要求回数３９が最大の論理範囲と、全アレイグループ１４１中、アレイグループ１４１毎の入出力要求回数３９の総和が最小のアレイグループ１４１内の入出力要求回数３９が最小の論理範囲の交換を実施すると決定できる。但し、前記（Ｄ）の移動処理の最後では、新配置情報テーブルは移動した分割領域のみの情報が定義、記憶されているので、その他の分割領域の情報は、旧配置情報テーブルよりコピーしておく。

【００４２】

また、本実施の形態では、制御装置１２の配下に、空き領域を含むアレイグループ１４１が在ることを条件としているが、この条件に関係なく、キャッシュメモリ１２３に仮想のアレイグループ１４１を定義する（すなわち、制御情報用メモリ１２２上の配置情報テーブル４０内の位置情報をキャッシュメモリ１２３上のアドレスにする）ことで作業領域は確保でき、同様の処理が可能となる。

【００４３】

以上の説明から明らかなように、本実施の形態の外部記憶システムの論理デバイスデータ交換方法によれば、上位の処理装置１０からの入出力要求の頻度が高い論理デバイス１４３もしくは論理デバイス１４３の一部を、他のアレイグループ１４１に移動し、特定のアレイグループ１４１内の特定の物理デバイス１４２に入出力負荷が集中することを回避して負荷を軽減させ、かつ複数のアレイグループ１４１間の負荷を均等とすることにより、処理装置１０からの入出力要求を高速に処理することが可能となる。換言すれば、アレイ型外部記憶装置１４を構成するすべての物理デバイス１４２の可用性が向上する。

【００４４】

また、業務に先立ち、利用者が高速入出力処理を期待するファイルに対応する論理デバイス１４３を、制御装置１２が備えるサービス機器１３を用いて指定するだけで、任意のアレイグループ１４１内の物理デバイス１４２の使用負荷を低減させかつアレイグループ１４１間の負荷を均等にすることにより、処理装置１０からの入出力要求を高速に処理することが可能である。

【００４５】

また、業務形態等の必要に応じて、たとえば入出力要求回数の大小等に関係なく、特定の論理デバイス１４３を特定のアレイグループ１４１に割り当てる等の多様かつ柔軟な設定が可能となる。

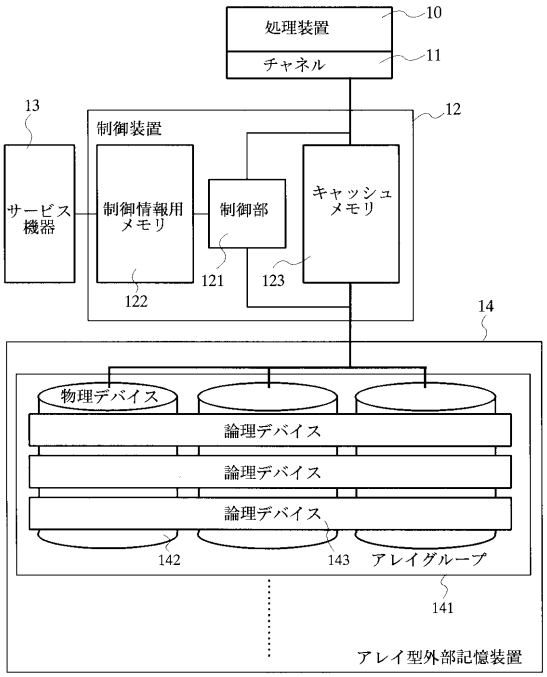
【００４６】

以上本発明者によってなされた発明を実施の形態に基づき具体的に説明したが、本発明は前記実施の形態に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【００４７】
【発明の効果】
本発明によれば、物理デバイス上に複数の論理デバイスを設定して稼動する外部記憶システムを含む情報処理システムにおける論理デバイスの入出力性能の向上を実現することができる、という効果が得られる。
【００４８】
また、本発明によれば、物理デバイス上に複数の論理デバイスを設定して稼動する外部記憶システムを含む情報処理システムにおける物理デバイスの可用性の向上を実現することができる、という効果が得られる。
【００４９】
また、本発明によれば、物理デバイス上に複数の論理デバイスを設定して稼動する外部記憶システムを含む情報処理システムにおいて、利用者の要請に合わせた論理デバイスの設定の効率化および多様化を実現することができる、という効果が得られる。
【図面の簡単な説明】
【図１】本発明の一実施の形態である外部記憶システムの論理デバイスデータ交換方法が適用される情報処理システムの構成の一例を示す概念図である。
【図２】本発明の一実施の形態である外部記憶システムの論理デバイスデータ交換方法を実施する情報処理システムの構成の一部を取り出して示す概念図である。
【図３】本発明の一実施の形態である外部記憶システムの論理デバイスデータ交換方法にて用いられる制御情報の一例を示す概念図である。
【図４】本発明の一実施の形態である外部記憶システムの論理デバイスデータ交換方法の作用の一例を示すフローチャートである。
【図５】本発明の一実施の形態である外部記憶システムの論理デバイスデータ交換方法の作用の一例を示すフローチャートである。
【図６】本発明の一実施の形態である外部記憶システムの論理デバイスデータ交換方法の作用の一例を示すフローチャートである。
【図７】本発明の一実施の形態である外部記憶システムの論理デバイスデータ交換方法を実施する情報処理システムの仕様の一例を示す説明図である。
【図８】本発明の一実施の形態である外部記憶システムの論理デバイスデータ交換方法を実施する情報処理システムにおける制御情報の画面表示出力の一例を示す概念図である。
【符号の説明】
１０…処理装置（外部処理装置）、１１…チャネル、１２…制御装置、１３…サービス機器、１３ａ…操作画面、１３ｂ…ボックス、１３ｃ…棒グラフ、１４…アレイ型外部記憶装置、２０…分割領域、３０…交換対象論理デバイス番号登録テーブル、３１…開始時刻テーブル、３２…自動交換機能有効フラグ、３３…論理デバイス管理テーブル、３４…論理デバイス番号、３５…論理デバイス容量、３６…アレイグループ番号、３７…論理デバイス数、３８…アレイグループ間変換テーブル、３８ａ…新配置情報テーブルポインタ、３８ｂ…旧配置情報テーブルポインタ、３９…入出力要求回数、４０…配置情報テーブル、４１…接続アレイグループ番号テーブル、４２…物理デバイス容量、１２１…制御部、１２２…制御情報用メモリ、１２３…キャッシュメモリ、１４１…アレイグループ、１４２…物理デバイス、１４３…論理デバイス。

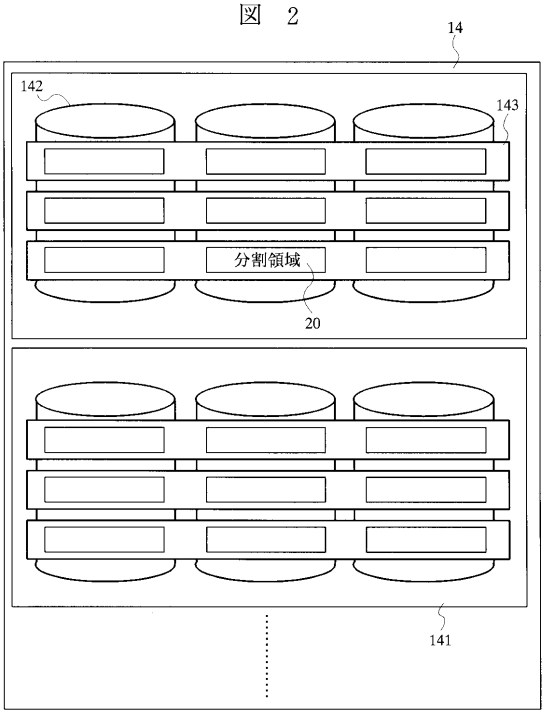
【図１】

図 1



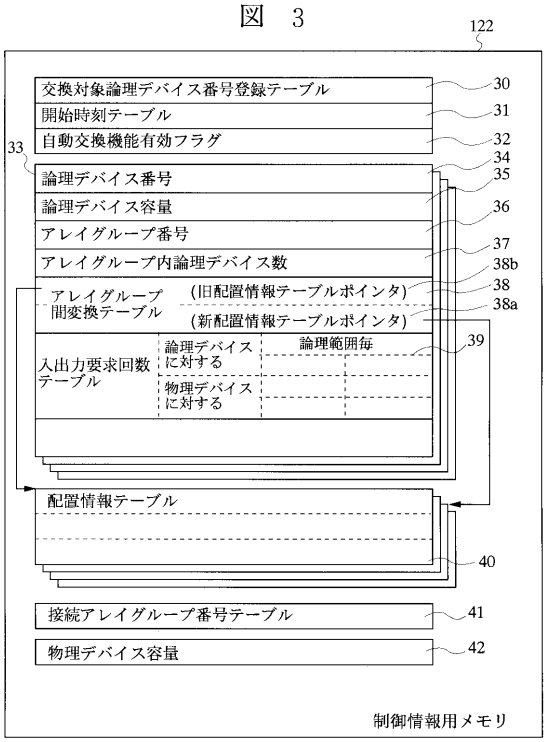
【図２】

図 2



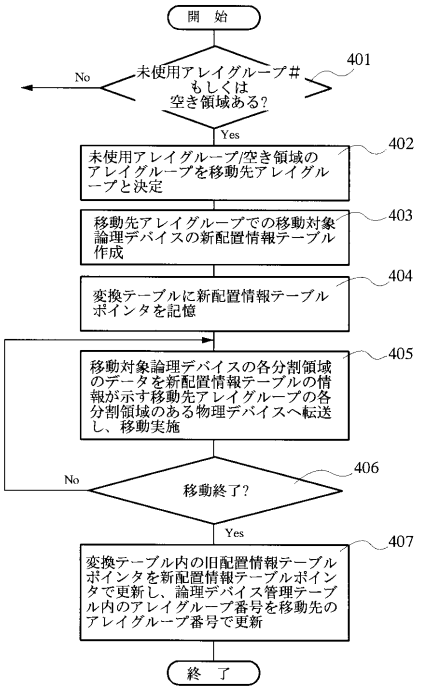
【図３】

図 3



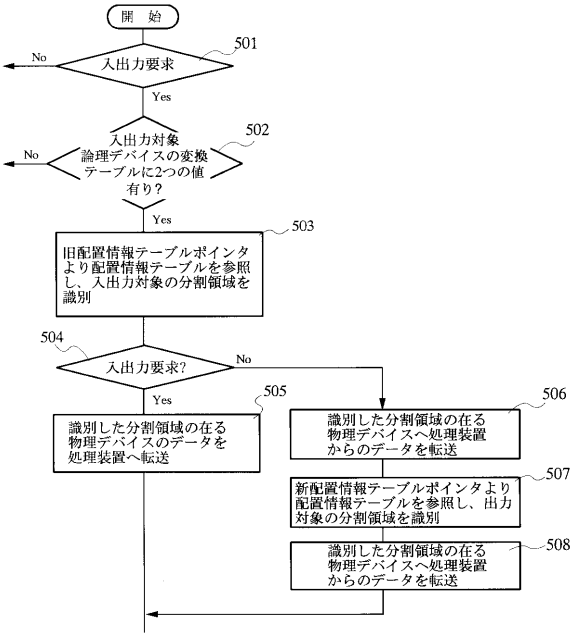
【図４】

図 4



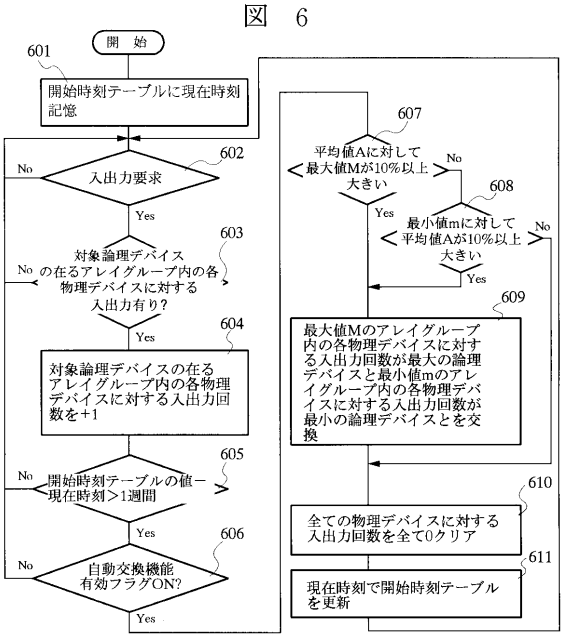
【図５】

図 5



【図６】

図 6



・アレイグループ内の物理デバイスに対する入出力回数の最大値を最大値Mとする。
・アレイグループ内の物理デバイスに対する入出力回数の最小値を最小値mとする。
・アレイグループ毎の物理デバイスに対する入出力回数の制御装置下の平均値を平均値Aとする。

【図７】

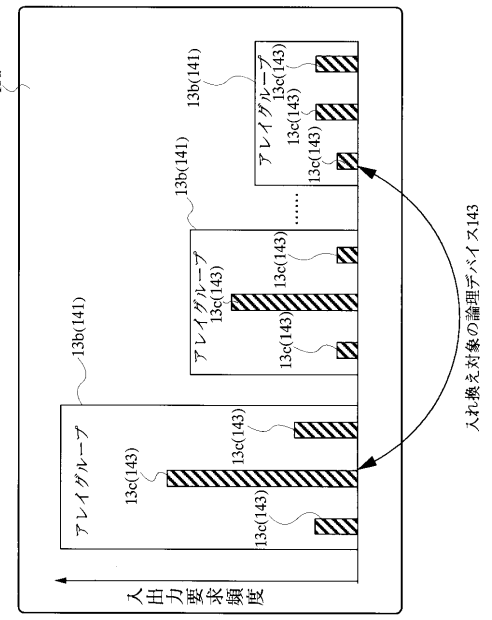
図 7

論理デバイスの移動・交換の一実施例の仕様一覧表

機能	項目			仕様
移動	移動元アレイグループ	物理デバイス数		1台以上
		論理デバイス数		1台以上
	移動先アレイグループ	物理デバイス数		1台以上
		論理デバイス数		1台以上
	移動の単位			キャッシュメモリを使用すれば移動先アレイグループは不要
				論理デバイス
				論理デバイス内のある一部(論理範囲)
	ホストからの入出力動作	移動中	入力	移動元アレイグループ内のデータを対象
			出力	移動元と移動先のアレイグループ内の両データを対象
		移動後	入力	移動先アレイグループ内のデータを対象
			出力	移動先アレイグループ内のデータを対象
交換	交換条件			上記の移動機能が動作できること
	交換契機			自動もしくは外部より指示

【図８】

図 8



フロントページの続き

(72)発明者　川口　勝洋
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
(72)発明者　中西　弘晃
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０７－０５６６９１（ＪＰ，Ａ）
特開平０９－０６９０３１（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.[7]，ＤＢ名)
G06F　3/06-08
G06F 12/00