

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4703738号
(P4703738)

(45) 発行日 平成23年6月15日(2011.6.15)

(24) 登録日 平成23年3月18日(2011.3.18)

(51) Int.Cl. F I
G06F 3/06 (2006.01) G O 6 F 3/06 3 O 4 F
G06F 13/10 (2006.01) G O 6 F 13/10 3 4 O B

請求項の数 5 (全 15 頁)

(21) 出願番号 特願2009-65340 (P2009-65340)
 (22) 出願日 平成21年3月18日(2009.3.18)
 (65) 公開番号 特開2010-218320 (P2010-218320A)
 (43) 公開日 平成22年9月30日(2010.9.30)
 審査請求日 平成22年4月16日(2010.4.16)

(73) 特許権者 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 100070150
 弁理士 伊東 忠彦
 (74) 代理人 100146776
 弁理士 山口 昭則
 (72) 発明者 大海 成浩
 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
 審査官 坂東 博司

最終頁に続く

(54) 【発明の名称】 ストレージ装置

(57) 【特許請求の範囲】

【請求項1】

非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能をサポートする分散キャッシュメモリ型のストレージ装置であって、

ネットワークを介して接続された複数の記憶装置を備え、

コピー元の記憶装置は、受領した一括バッファ新規設定コマンドに含まれるパラメータをチェックし、一括バッファ新規設定する記憶装置のグループ及びコピー先の記憶装置を確認して前記コピー先の指定された記憶装置と通信を行って前記パラメータを前記指定された記憶装置へ送信する手段を有し、

前記指定された記憶装置は、前記コピー元の記憶装置からの前記パラメータに基づいて前記指定された記憶装置内のバッファ新規設定処理を行って前記コピー元の記憶装置へ前記バッファ新規設定処理の結果を通知する手段を有し、

前記パラメータは、前記グループ数、前記コピー先の記憶装置を識別する装置識別子、バッファサイズ、及び指定する記憶装置がコピー元であるかコピー先であることを示すバッファ使用用途を含む、ストレージ装置。

【請求項2】

前記一括バッファ新規設定コマンドは、前記コピー元の記憶装置内、或いは、前記コピー元の記憶装置に外部接続されたコンピュータから発行される、請求項1記載のストレージ装置。

【請求項3】

10

20

前記コピー元の記憶装置は、前記指定した記憶装置との通信の成否を判断し、前記パラメータの少なくとも一部を格納したバッファテーブルに判断結果を示す通信成否判断フラグを格納する手段を有する、請求項 1 又は 2 記載のストレージ装置。

【請求項 4】

前記コピー元の記憶装置は、複数のグループの一括バッファ新規設定処理を行う手段を有する、請求項 1 乃至 3 のいずれか 1 項記載のストレージ装置。

【請求項 5】

前記コピー元の記憶装置は、前記指定された記憶装置との間の通信が経路の閉塞により失敗した後、経路の復旧した時に一括バッファ新規設定を自動的に継続させる手段を有する、請求項 1 乃至 4 のいずれか 1 項記載のストレージ装置。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ装置に係り、特にリモート・イクイバレント・コピー（REC：Remote Equivalent Copy）モード（Mode）等のリモートコピー機能を有するストレージ装置に関する。

【背景技術】

【0002】

R A I D（Redundant Array of Independent Disks）装置間のバックアップやミラーリングを行うため、アドバンスド・コピー（Advanced Copy）において順序性を保障した R E C コンシステンシ・モード（Consistency Mode）機能をサポートする分散キャッシュメモリ型のストレージ装置が提案されている。アドバンスド・コピーとは、ストレージ装置が任意の時点でデータ（ソースボリューム）をコピーして別のボリューム（コピーボリューム）を同じストレージ装置内で作成する機能を言う。

20

【0003】

リモートミラーリングには、同期モードと非同期モードがある。同期モードでは、書き込み要求に対する完了報告はデータのコピー先への書き込みが完了するまで行われない。一方、非同期モードでは、書き込み要求に対する完了報告はデータのコピー先への書き込みの完了を待たずに行われる。非同期モードには、リモートミラーリング用のコンシステンシ・モードと、リモートバックアップ用のスタックモードがあり、コンシステンシ・モードでは書き込み順序性が保証される。

30

【0004】

R E C モード自体は周知であり、例えば開始（Start）、停止（Stop）、中止（Suspend）、強制停止（Forced Stop）、強制中止（Forced Suspend）、再開（Resume）等のコマンドに従ったミラーリングを行う。開始コマンドは、ミラーリングされたボリュームの作成と同期処理の開始を指示する。停止コマンドは、ミラーリングの解除、ミラーリングされたボリュームの破棄、この停止コマンドが発行される以前に書き込まれた全てのデータをこの停止コマンドが終了するまでコピーボリュームに転送することを指示する。中止コマンドは、ミラーリングを解除するがミラーリングされたボリュームは保持し、この中止コマンドが発行される以前に書き込まれた全てのデータをこの停止コマンドが終了するまでコピーボリュームに転送することを指示するものであり、保持されたミラーリングされたボリュームは再開コマンドにより再同期可能である。強制停止コマンドは、ミラーリングの解除、ミラーリングされたボリュームの破棄を指示し、この強制停止コマンドが発行される以前に書き込まれたデータはコピーボリュームに転送する必要はない。強制中止コマンドは、ミラーリングを解除するがミラーリングされたボリュームは保持し、この強制中止コマンドが発行される以前に書き込まれたデータはコピーボリュームに転送する必要はなく、一方、保持されたミラーリングされたボリュームは再開コマンドにより再同期可能である。再開コマンドは、中止コマンドの後の再同期を指示するものであり、再同期は中止状態で書き込まれたデータを転送することで実行される。

40

【0005】

50

図 1 は、従来の分散キャッシュメモリ型のストレージ装置の一例を説明する図である。図 1 において、ストレージ装置 1 は、ネットワーク 2 を介して接続されたコピー元の記憶装置（以下、コピー元装置と言う）3 - 1 及びコピー先の記憶装置（以下、コピー先装置と言う）3 - 2 を有する。コピー元装置 3 - 1 及びコピー先装置 3 - 2 は、例えば複数の制御モジュール 1 1 を備えた同じ構成を有する。各制御モジュール 1 1 は、中央処理装置（CPU：Central Processing Unit）等のプロセッサ、記録用バッファ 1 2 及びディスク装置 1 3 を有する。

【0006】

図 1 に示す如きストレージ装置 1 では、論理ボリュームのデータのリード及び/又はライト（以下、単にリード/ライトと言う）処理を制御モジュール 1 1 毎に実行している。又、例えばアドバンスト・コピーでの順序性を保障した REC コンシステンシ・モード機能のように、非同期モードでリモートコピーを行う際の書き込み順序性が保証された場合は、コピー元装置 3 - 1 内においてディスク装置 1 3 のディスクに格納されたデータをコピー元装置 3 - 1 内のバッファ 1 2 に格納し、バッファ 1 2 がフルになるか、若しくは、データの格納開始から一定時間経過すると、コピー元装置 3 - 1 のバッファ 1 2 内のデータをまとめてネットワーク 2 を介して接続されたコピー先装置 3 - 2 へ送信する。コピー先装置 3 - 2 では、データの受信を全て完了すると、受信して記録用バッファ 1 2 内に格納されたデータをまとめてコピー先のディスク装置 1 3 のディスクへ展開して格納する。このように、REC コンシステンシ・モード機能を実行するには、コピー元装置 3 - 1 及びコピー先装置 3 - 2 の夫々において、バッファ 1 2 の設定が必要となる。

【0007】

図 2 は、REC コンシステンシ・モードで用いられるバッファ 1 2 のバッファ領域新規設定手順の一例を説明する図である。図 2 中、(a) は REC コンシステンシ・モード機能を実行するために一例として 3 台の記憶装置 3 - 1 ~ 3 - 3 を用意した状態、(b) は各記憶装置 3 - 1 ~ 3 - 3 をケーブル又はネットワーク 5 で接続した状態を示す。ここでは説明の便宜上、記憶装置 3 - 1, 3 - 2 がグループ G 1 を形成し、記憶装置 3 - 1, 3 - 3 がグループ G 2 を形成するものとする。図 2 中、(c) は各記憶装置 3 - 1 ~ 3 - 3 にどの記憶装置と接続するべきかを各記憶装置 3 - 1 ~ 3 - 3 のストレージ管理パーソナルコンピュータ（PC：Personal Computer）6 から設定する経路設定を行う状態を示す。図 2 中、(d) はグループ G 1 に対して REC コンシステンシ・モード機能を実行するためのバッファ 1 2 の新規設定を、記憶装置 3 - 1, 3 - 2 のストレージ管理 PC 6 から行う状態を示し、図 2 (e) はグループ G 2 に対して REC コンシステンシ・モード機能を実行するためのバッファ 1 2 の新規設定を、記憶装置 3 - 1, 3 - 3 のストレージ管理 PC 6 から行う状態を示す。

【0008】

このように、バッファ 1 2 の新規設定は、コピー元装置 3 - 1 及びコピー先装置 3 - 2, 3 - 3 の夫々で対応するストレージ管理 PC 6 から行う必要がある。このため、バッファ 1 2 の新規設定をするには、コピー元装置 3 - 1 及びコピー先装置 3 - 2, 3 - 3 の夫々において新規設定を行う必要がある。又、コピー元装置 3 - 1 及びコピー先装置 3 - 2, 3 - 3 の夫々でストレージ管理 PC 6 が必要となる。

【0009】

更に、コピー元装置 3 - 1 及びコピー先装置 3 - 2, 3 - 3 の各記憶装置毎の新規設定になり、相手側記憶装置のバッファ 1 2 が正常に新規設定されているかを知るためにも、コピー元装置 3 - 1 及びコピー先装置 3 - 2, 3 - 3 の夫々のストレージ管理 PC 6 からの確認が必要となる。つまり、コピー元装置 3 - 1 のストレージ管理 PC 6 は、相手側記憶装置である各コピー先装置 3 - 2, 3 - 3 のバッファ 1 2 が正常に新規設定されているか否かを確認する必要がある。各コピー先装置 3 - 2, 3 - 3 のストレージ管理 PC 6 は、相手側記憶装置であるコピー元装置 3 - 1 のバッファ 1 2 が正常に新規設定されているか否かを確認する必要がある。

【先行技術文献】

10

20

30

40

50

【特許文献】

【0010】

【特許文献1】特開2005-122235号公報

【発明の概要】

【発明が解決しようとする課題】

【0011】

従来のストレージ装置では、例えばRECコンシステンシ・モードのように非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能で用いられるバッファの新規設定を簡単に行うことはできないという問題があった。

【0012】

そこで、本発明は、非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能で用いられるバッファの新規設定を簡単に行うことのできる記憶装置を提供することを目的とする。

【課題を解決するための手段】

【0013】

本発明の一観点によれば、非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能をサポートする分散キャッシュメモリ型のストレージ装置であって、ネットワークを介して接続された複数の記憶装置を備え、コピー元の記憶装置は、受領した一括バッファ新規設定コマンドに含まれるパラメータをチェックし、一括バッファ新規設定する記憶装置のグループ及びコピー先の記憶装置を確認して前記コピー先の指定された記憶装置と通信を行って前記パラメータを前記指定された記憶装置へ送信する手段を有し、前記指定された記憶装置は、前記コピー元の記憶装置からの前記パラメータに基づいて前記指定された記憶装置内のバッファ新規設定処理を行って前記コピー元の記憶装置へ前記一括バッファ新規設定処理の結果を通知する手段を有し、前記パラメータは、前記グループ数、前記コピー先の記憶装置を識別する装置識別子、バッファサイズ、及び指定する記憶装置がコピー元であるかコピー先であることを示すバッファ使用用途を含むストレージ装置が提供される。

【発明の効果】

【0014】

開示の記憶装置によれば、非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能で用いられるバッファの新規設定を簡単に行うことができる。

【図面の簡単な説明】

【0015】

【図1】従来の分散キャッシュメモリ型のストレージ装置の一例を説明する図である。

【図2】RECコンシステンシ・モードで用いられるバッファのバッファ領域新規設定手順の一例を説明する図である。

【図3】本発明の一実施例において、分散キャッシュメモリ型のストレージ装置のRECコンシステンシ・モードで用いられるバッファのバッファ領域新規設定手順の一例を説明する図である。

【図4】REC一括バッファ新規設定コマンドのパラメータの一例を示す図である。

【図5】パラメータチェック処理を説明するフローチャートである。

【図6】通信処理を説明するフローチャートである。

【図7】ステップS25で取得されるパラメータの一例を示す図である。

【図8】処理ST3を説明するフローチャートである。

【図9】処理ST4を説明するフローチャートである。

【図10】RECバッファテーブルの一例を示す図である。

【図11】処理ST5を説明するフローチャートである。

【図12】処理ST6を説明するフローチャートである。

【図13】ストレージ装置のハードウェア構成の一例を示すブロック図である。

【発明を実施するための形態】

10

20

30

40

50

【 0 0 1 6 】

開示のストレージ装置は、例えばリモート・イクイバレント・コピー（REC）コンシステンシ・モード機能のように非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能をサポートする分散キャッシュメモリ型であり、ネットワークを介して接続された複数の記憶装置を備える。コピー元の記憶装置は、受領した一括バッファ新規設定コマンドに含まれるパラメータをチェックし、一括バッファ新規設定する記憶装置のグループ及びコピー先の記憶装置を確認して指定された記憶装置と通信を行ってパラメータを指定された記憶装置へ送信する手段を有する。一方、指定された記憶装置は、コピー元の記憶装置からのパラメータに基づいて指定された記憶装置内のバッファ新規設定処理を行ってコピー元の記憶装置へバッファ新規設定処理の結果を通知する手段を有する。パラメータは、グループ数、コピー先の記憶装置を識別する装置識別子、バッファサイズ、及び指定する記憶装置がコピー元であるかコピー先であることを示すバッファ使用用途を含む。

10

【 0 0 1 7 】

つまり、例えばRECコンシステンシ・モードのように非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能を実現させるのに必要なバッファの設定を、記憶装置間を繋ぐ経路を利用して一括新規設定をする。このようにして、RECコンシステンシ・モード機能のように非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能に対する拡張機能を提供可能となる。

20

【 0 0 1 8 】

以下に、本発明のストレージ装置の各実施例を、図3以降と共に説明する。

【実施例】

【 0 0 1 9 】

図3は、本発明の一実施例において、分散キャッシュメモリ型のストレージ装置のRECコンシステンシ・モードで用いられる記憶装置内の記録用バッファのバッファ領域新規設定手順の一例を説明する図である。ストレージ装置21は、ネットワークを介して接続されたコピー元の記憶装置（以下、コピー元装置と言う）23-1及び又は複数のコピー先の記憶装置（以下、コピー先装置と言う）23-2、23-3を有する。コピー元装置23-1及びコピー先装置23-2、23-3の各々は、例えば図1の場合と同様に複数の制御モジュールを備えた同じ構成を有しても、或いは、少なくとも1つの制御モジュールを備えたものであっても良い。制御モジュールは、CPU等のプロセッサ、記録用バッファ及びディスク装置を有する。尚、各記憶装置23-1～23-3内に設けられる制御モジュールの数は同じである必要はない。

30

【 0 0 2 0 】

図3中、(a)はRECコンシステンシ・モード機能を実行するために一例として3台の記憶装置23-1～23-3を用意した状態、(b)は各記憶装置23-1～23-3をケーブル又はネットワーク25で接続した状態を示す。ここでは説明の便宜上、記憶装置23-1、23-2がグループG1を形成し、記憶装置23-1、23-3がグループG2を形成するものとする。図3中、(c)は各記憶装置23-1～23-3にどの記憶装置と接続するべきかを各記憶装置23-1～23-3のストレージ管理パーソナルコンピュータ(PC: Personal Computer)26、或いは、図示を省略したホストサーバ（以下、単にストレージ管理PC26と言う）から設定する経路設定を行う状態を示す。尚、ストレージ管理PC26とホストサーバは、共に設けられていても良い。図3中、(d)は複数のグループG1、G2に対してRECコンシステンシ・モード機能を実行するためのバッファの新規設定を、記憶装置23-1のストレージ管理PC26から一括で行う状態を示す。

40

【 0 0 2 1 】

このように、バッファの新規設定は、コピー元装置23-1の筐体のストレージ管理PC26から一又は複数のコピー先装置23-2、23-3に対して一括で行えば良い。このため、バッファの新規設定をするのに、コピー先装置23-2、23-3の夫々におい

50

て新規設定を行う必要はない。又、バッファの新規設定のために、コピー先装置 23 - 2 , 23 - 3 でストレージ管理 PC 26 を必要とすることはない。

【 0 0 2 2 】

更に、相手側記憶装置のバッファが正常に新規設定されているかを知るために、コピー元装置 3 - 1 のストレージ管理 PC 26 から確認する必要はない。

【 0 0 2 3 】

つまり、バッファの新規設定を行う時に記憶装置間において設定済みである経路、即ち、ケーブル又はネットワーク 25 を経由して、相手側記憶装置と通信を行い、相手側記憶装置のバッファの新規設定を一括で行うようにする。又、相手側記憶装置と通信を行い、この通信に対する応答に基づいて相手側記憶装置のバッファの新規設定を確認することが可能となる。

10

【 0 0 2 4 】

相手側記憶装置へ新規設定のための通信を行い、記憶装置間の経路の切断や相手側記憶装置の電源断等で通信が失敗して相手側記憶装置への新規設定が失敗した際は、自記憶装置において通信が失敗したことを記憶し、相手側記憶装置への経路が再度繋がった際に自動的に相手側記憶装置へ通信を行ってバッファの新規設定を行う。このように、一方の記憶装置に対する新規設定コマンドで、元先両記憶装置のバッファが一括で新規設定できる。

【 0 0 2 5 】

次に、本実施例において各記憶装置の制御モジュール内の CPU が実行する処理 ST 1 ~ ST 6 について説明する。処理 ST 1 ~ ST 6 を実行する CPU は、後述する図 13 に示す CPU 15 に相当し、各処理 ST 1 ~ ST 6 を実行する各手段（又は、機能部）を形成する。処理 ST 1 は、記憶装置においてストレージ管理 PC 6 より一括でバッファ新規設定を行うための命令（以下、REC 一括バッファ新規設定コマンドと言う）を受領する処理である。処理 ST 2 は、記憶装置において受領した REC 一括バッファ新規設定コマンドのパラメータをチェックし、新規設定するグループ及び相手側記憶装置を確認して指定された相手側記憶装置と通信を行う処理である。処理 ST 3 は、指定された相手側記憶装置で REC 一括バッファ新規設定コマンドのパラメータに基づいた一括バッファ新規設定を行う処理である。処理 ST 4 は、記憶装置において指定した相手側記憶装置との通信の成否を判断する処理である。処理 ST 5 は、REC 一括バッファ新規設定コマンドを受領した記憶装置側で複数のグループの一括バッファ新規設定を行う処理である。処理 ST 6 は、記憶装置間の通信が失敗した後、経路復旧した時にバッファ新規設定を自動的に継続させる処理である。

20

30

【 0 0 2 6 】

(1) 処理 ST 1 :

処理 ST 1 では、通信元（例えば、コピー元）の記憶装置がストレージ管理 PC 26 より一括バッファ新規設定を行うための REC 一括バッファ新規設定コマンドを受領する。処理 ST 1 が終了すると、処理は後述する処理 ST 2 へ進む。

【 0 0 2 7 】

図 4 は、REC 一括バッファ新規設定コマンドのパラメータの一例を示す図である。図 4 に示すように、パラメータには、新規設定グループ数 SGN、相手装置識別子 SI 1 ~ SI n (n は 2 以上の自然数)、REC バッファサイズ BS 1 ~ BS n 及び REC バッファ使用用途 BU 1 ~ BU n が含まれる。新規設定グループ数 SGN は、受領した REC 一括バッファ新規設定コマンドで一括バッファ新規設定する記憶装置のグループの数を示す。相手装置識別子 SI 1 ~ SI n は、一括バッファ新規設定を行う対となる記憶装置を識別する識別子を示し、記憶装置内のバッファやディスク装置等の記憶部に設定されている経路情報より取得する。経路情報は、上記の経路設定を示す情報である。REC バッファサイズ BS 1 ~ BS n は、使用するバッファサイズを示す。REC バッファ使用用途 BU 1 ~ BU n は、指定する記憶装置が送信側（即ち、通信元又はコピー元）であるか受信側（即ち、通信先又はコピー先）であることを示す。

40

50

【 0 0 2 8 】

(2) 処理 S T 2 :

処理 S T 2 では、通信元の記憶装置が処理 S T 1 で受領した R E C 一括バッファ新規設定コマンドのパラメータをパラメータチェック処理でチェックし、一括バッファ新規設定する記憶装置のグループ及び相手側記憶装置を確認して指定された相手側記憶装置との通信を通信処理で行う。処理 S T 2 は、グループ内の各相手側記憶装置に対して行われる。

【 0 0 2 9 】

図 5 は、パラメータチェック処理を説明するフローチャートである。図 5 において、ステップ S 2 1 では、通信元の記憶装置が処理 S T 1 で受領した R E C 一括バッファ新規設定コマンドのパラメータから、指定された相手側記憶装置（例えば、コピー先の記憶装置）の装置識別子 S I が R E C 一括バッファ新規設定コマンドを受領した記憶装置内の記憶部に設定されている経路情報に存在するか否かを判定する。ステップ S 2 1 の判定結果が Y E S であると、ステップ S 2 2 では、指定された相手側記憶装置において R E C バッファサイズ B S で指定されたバッファサイズを新規設定可能であるか否かを判定する。ステップ S 2 2 の判定結果が Y E S であれば、処理は終了して図 6 と共に後述する通信処理へ進む。一方、ステップ S 2 1 又はステップ S 2 2 の判定結果が N O であると、ステップ S 2 3 では R E C 一括バッファ新規設定コマンドの実行が失敗したと判断し、指定された相手側記憶装置の経路情報がない（ステップ S 2 1 で N O ）又は指定されたバッファサイズのバッファ獲得が失敗した（ステップ S 2 2 で N O ）ことを示すコマンド失敗を R E C 一括バッファ新規設定コマンドの発行元である通信元の記憶装置のストレージ管理 P C 2 6 へ通知する。

【 0 0 3 0 】

図 6 は、通信処理を説明するフローチャートである。図 6 において、ステップ S 2 5 では、通信元の記憶装置が R E C バッファ一括新規設定コマンドで指定されたパラメータ及び R E C 一括バッファ新規設定コマンドを受領した記憶装置内の記憶部に設定されている経路情報に基づいて、各相手側記憶装置に送信すべきパラメータを作成する。

【 0 0 3 1 】

図 7 は、ステップ S 2 5 で取得されるパラメータの一例を示す図である。図 7 に示すように、取得されるパラメータには、通信元の記憶装置の装置識別子（以下、通信元装置識別子という）S I s、R E C バッファサイズ B S 及び R E C バッファ使用用途 B U が含まれる。通信元装置識別子 S I s は、通信元の記憶装置を識別する識別子を設定するものであり、R E C 一括バッファ新規設定コマンドを受領した記憶装置内の記憶部に設定されている経路情報に基づいて設定する。R E C バッファサイズ B S は、使用するバッファサイズを設定する。R E C バッファ使用用途 B U は、指定する相手側記憶装置が送信側であるか受信側であるかを設定する。

【 0 0 3 2 】

ステップ S 2 6 では、取得したパラメータを相手側記憶装置へ送信する。ステップ S 2 7 では、一括バッファ新規設定する記憶装置のグループ内の全ての記憶装置へのパラメータの送信が完了したか否かを判定する。ステップ S 2 7 の判定結果が N O であると処理はステップ S 2 5 へ戻り、Y E S であると処理は終了して後述する処理 S T 3 へ進む。

【 0 0 3 3 】

(3) 処理 S T 3 :

処理 S T 3 では、指定された相手側記憶装置、即ち、通信先（例えば、コピー先）の記憶装置において、指定されたパラメータによるバッファ新規設定処理を行う。処理 S T 3 は、グループ内の各相手側記憶装置において行われる。

【 0 0 3 4 】

図 8 は、処理 S T 3 を説明するフローチャートである。図 8 において、ステップ S 3 1 では、指定された相手側記憶装置において、R E C バッファサイズ B S で指定されたバッファサイズを新規設定可能であるか否かを判定する。ステップ S 3 1 の判定結果が Y E S であると、ステップ S 3 2 では、指定された相手側記憶装置において、R E C バッファサ

10

20

30

40

50

イズBSで指定されたバッファサイズのバッファを獲得し、獲得したバッファに関する情報を含むRECバッファテーブルを作成して指定された相手側記憶装置内の記憶部に格納すると共に、通信元の記憶装置へ処理ST3の正常終了を通知する。RECバッファテーブルについては後述する。ステップS32の後、処理は終了して図9と共に後述する処理ST4へ進む。

【0035】

一方、ステップS31の判定結果がNOであると、ステップS33では、指定された相手側記憶装置において、REC一括バッファ新規設定コマンドの実行が失敗したと判断し、指定されたバッファサイズのバッファの獲得が失敗したことを示すコマンド失敗をREC一括バッファ新規設定コマンドの発行元、即ち、通信元の記憶装置へ通知することで通信元の記憶装置のストレージ管理PC26へ通知する。

10

【0036】

(4)処理ST4:

処理ST4では、通信元の記憶装置と指定した相手側記憶装置(又は、通信先の記憶装置)との通信の成否を通信成否判断処理で判断する。処理ST4は、グループ内の各相手側記憶装置に対して行われる。

【0037】

図9は、処理ST4を説明するフローチャートである。図9において、ステップS41では、通信元の記憶装置において、相手側記憶装置の一括バッファ新規設定が成功したか否かを判定する。コマンド失敗の通知がなくステップS41の判定結果がYESであると、ステップS42では、通信元の記憶装置において、相手側記憶装置への通信が成功したか否かを判定する。相手側記憶装置から処理ST3の正常終了の通知がなくステップS42の判定結果がNOであると、ステップS43では、RECバッファテーブルに通信失敗を示す通信成否判断フラグCJFを立てて、処理はステップS44へ進む。

20

【0038】

相手側記憶装置から処理ST3の正常終了の通知がありステップS42の判定結果がYESであるか、或いは、ステップS43の後、ステップS44では、一括バッファ新規設定する記憶装置のグループ内の全ての記憶装置との通信が完了したか否かを判定する。ステップS44の判定結果がNOであると処理はステップS41へ戻り、YESであると処理は終了して後述する処理ST5へ進む。

30

【0039】

一方、指定されたバッファサイズのバッファ獲得が失敗したことを示すコマンド失敗が相手側記憶装置から通知されてステップS41の判定結果がNOであると、ステップS45ではREC一括バッファ新規設定コマンドの実行が失敗したと判断し、コマンド失敗をREC一括バッファ新規設定コマンドの発行元である通信元の記憶装置のストレージ管理PC26へ通知する。

【0040】

図10は、RECバッファテーブルの一例を示す図である。図10に示すように、RECバッファテーブルには、相手装置識別子SI1~SIn、RECバッファサイズBS1~BSn、RECバッファ使用用途BU1~BUn及び通信成否判断フラグCFJ1~CFJnが含まれる。尚、本実施例では、相手側記憶装置から正常終了の通知がないと通信失敗を示す通信成否判断フラグCJFを立てているが、相手側記憶装置から正常終了の通知があると通信成功を示す通信成否フラグCFJを立てるようにしても良いことは言うまでもない。要は、通信成否フラグCFJで通信の成否を示すことができれば良い。

40

【0041】

(5)処理ST5:

処理ST5は、REC一括バッファ新規設定コマンドを受領した通信元の記憶装置で複数のグループの一括バッファ新規設定処理を行う。処理ST5は、グループ内の各相手側記憶装置に対して行われる。

【0042】

50

図11は、処理ST5を説明するフローチャートである。図11において、ステップS51では、通信元の記憶装置において、各相手側記憶装置から獲得された、指定されたバッファサイズのバッファに関する情報を新規設定グループ数SGN分のRECバッファテーブルに格納することで一括バッファ新規設定を行う。各新規設定グループに対応するRECバッファテーブルに格納される、指定されたバッファサイズのバッファに関する情報には、相手装置識別子SI1～SIn、RECバッファサイズBS1～BSn、RECバッファ使用用途BU1～BUn及び通信成否判断フラグCFJ1～CFJnが含まれる。ステップS52では、全てのグループ内の全ての記憶装置との通信が完了したか否かを判定する。ステップS52の判定結果がNOであると処理はステップS51へ戻り、YESであると処理はステップS53へ進む。

10

【0043】

ステップS53では、REC一括バッファ新規設定コマンドの発行元がストレージ管理PC26であるか否かを判定する。例えば、通信元の記憶装置にストレージ管理PC26が設けられており、且つ、この通信元の記憶装置にホストサーバも接続されている場合、REC一括バッファ新規設定コマンドの発行元がストレージ管理PC26であればステップS53の判定結果がYESとなり、発行元がホストサーバであればステップS53の判定結果がNOとなる。ステップS53の判定結果がNOであると、ステップS54では、一括バッファ新規設定の正常終了をホストサーバへ通知し、処理は終了する。一方、ステップS53の判定結果がYESであると、ステップS55では、一括バッファ新規設定の正常終了をストレージ管理PC26へ通知し、処理は終了する。

20

【0044】

(6)処理ST6:

処理ST6は、記憶装置間の通信が記憶装置間の経路の閉塞により失敗した後、経路の復旧した時に一括バッファ新規設定を自動的に継続させる経路開通処理を行うものであり、図6に示す通信処理の変形例に相当する。処理ST6のステップS61～S63、S65～S69は、グループ内の各相手側記憶装置に対して行われ、ステップS164はグループ内の各相手側記憶装置において行われる。ここでは説明の便宜上、通信元の記憶装置の相手側(通信先の)記憶装置に対する一括バッファ新規設定が通信元の記憶装置と相手側記憶装置との間の経路の閉塞により失敗すると、通信元の記憶装置内の記憶部に経路の閉塞を示す閉塞フラグが立てられるものとする。この閉塞フラグは、図10と共に説明したRECバッファテーブルに通信成否フラグCFJ1～CFJnと同様に設定されるものであっても良い。

30

【0045】

図12は、処理ST6を説明するフローチャートである。図12において、ステップS61では、通信元の記憶装置において、相手側記憶装置との通信がこの相手側記憶装置への経路の閉塞により失敗しているか否かを閉塞フラグに基づいて判定する。ステップS61の判定結果がYESであると、ステップS62では、図6のステップS25と同様にして相手側記憶装置に送信すべきパラメータを作成する。ステップS63では、図6のステップS26と同様に作成したパラメータを相手側記憶装置に送信する。

【0046】

ステップS164では、相手側記憶装置において図8に示す処理ST3と同様のバッファ新規設定処理が行われる。

40

【0047】

ステップS65では、通信元の記憶装置において、相手側記憶装置への送信の完了を確認し、ステップS66では、図9のステップS42と同様にして相手側記憶装置への通信が成功したか否かを判定する。ステップS66の判定結果がNOであると、ステップS67では、RECバッファテーブルに通信失敗を示す通信成否判断フラグCJFを立てて、処理はステップS68へ進む。ステップS68では、RECバッファテーブルに、通信成否判断フラグCJFを立てられた相手側記憶装置への経路に対する閉塞フラグを立てる経路閉塞処理を行う。

50

【 0 0 4 8 】

一方、ステップ S 6 1 の判定結果が N O、或いは、ステップ S 6 6 の判定結果が Y E S であると、処理はステップ S 6 9 へ進む。ステップ S 6 9 では、相手側記憶装置への経路に対する閉塞フラグを解除する経路開通処理を行う。

【 0 0 4 9 】

このようにして、処理 S T 3 で各グループの相手側記憶装置への通信が相手側記憶装置への経路の閉塞により失敗してその後経路が開通した際、相手側記憶装置に対する一括バッファ新規設定中に経路の閉塞によりバッファ新規設定が失敗していた場合には、経路開通処理の延長で相手側記憶装置に対する一括バッファ新規設定を再開させることができる。

10

【 0 0 5 0 】

図 1 3 は、ストレージ装置 2 1 のハードウェア構成の一例を示すブロック図である。図 1 3 中、図 1 及び図 3 と同一部分には同一符号を付し、その説明は省略する。図 1 3 の例では、コピー元の記憶装置 2 3 - 1 には単一の制御モジュール (C M : Control Module) 1 1 が設けられ、コピー先の記憶装置 2 3 - 2 には 2 つの C M 1 1 が設けられているが、上記の如く、C M 1 1 の数はこれに限定されるものではない。又、説明の便宜上、ストレージ管理 P C 2 6 は、記憶装置 2 3 - 1 に対するもののみを示す。更に、ホストサーバ 3 0 1 は、ストレージ装置 2 1 の一部を形成するものであっても、ストレージ装置 2 1 に外部接続されるものであっても良い。ホストサーバ 3 0 1 自体は、汎用コンピュータで形成可能である。

20

【 0 0 5 1 】

記憶装置 2 3 - 1 は、C M 1 1 及びディスク装置 1 3 に加え、ストレージ管理 P C 2 6 と記憶装置 2 3 - 1 との間のインタフェース制御部を形成する C G I (Common Gateway Interface) 2 3 1、記憶装置 2 3 - 1 とホストサーバ 3 0 1 との間のインタフェース制御部を形成する C A (Channel Adapter) 2 3 2、記憶装置 2 3 - 1 と記憶装置 2 3 - 2 との間のインタフェース制御部を形成する R A (Remote Adapter) 2 3 3 を有する。C M 1 1 は、バッファ (又は、キャッシュメモリ) 1 2 に加え、C P U 1 5 及び F C (Fiber Channel) 1 6 を有する。F C 1 6 は、C M 1 1 とディスク装置 1 3 との間のインタフェース制御部を形成する。尚、R E C バッファテーブルを格納するためのメモリ (図示せず) を、C M 1 1 内又は C M 1 1 外に設けても良い。

30

【 0 0 5 2 】

記憶装置 2 3 - 2 は、C M 1 1 が 2 つ設けられている点を除けば、基本的には記憶装置 2 3 - 1 と同様の構成を有する。ただし、図 1 3 では、バッファの新規設定のためにコピー先の記憶装置 2 3 - 2 でストレージ管理 P C 2 6 を必要とすることはないので、記憶装置 2 3 - 2 のストレージ管理 P C 2 6 の図示は省略してある。

【 0 0 5 3 】

ディスク装置 1 3 は、磁気ディスク装置に限定されず、光ディスク装置や光磁気ディスク装置であっても良い。又、ディスク装置 1 3 の代わりに各種記録媒体を用いる記録装置を使用可能であることは言うまでもない。

【 0 0 5 4 】

以上の実施例を含む実施形態に関し、更に以下の付記を開示する。

40

(付記 1)

非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能をサポートする分散キャッシュメモリ型のストレージ装置であって、

ネットワークを介して接続された複数の記憶装置を備え、

コピー元の記憶装置は、受領した一括バッファ新規設定コマンドに含まれるパラメータをチェックし、一括バッファ新規設定する記憶装置のグループ及びコピー先の記憶装置を確認して前記コピー先の指定された記憶装置と通信を行って前記パラメータを前記指定された記憶装置へ送信する手段を有し、

前記指定された記憶装置は、前記コピー元の記憶装置からの前記パラメータに基づいて

50

前記指定された記憶装置内のバッファ新規設定処理を行って前記コピー元の記憶装置へ前記バッファ新規設定処理の結果を通知する手段を有し、

前記パラメータは、前記グループ数、前記コピー先の記憶装置を識別する装置識別子、バッファサイズ、及び指定する記憶装置がコピー元であるかコピー先であることを示すバッファ使用用途を含む、ストレージ装置。

(付記 2)

前記一括バッファ新規設定コマンドは、前記コピー元の記憶装置内、或いは、前記コピー元の記憶装置に外部接続されたコンピュータから発行される、付記 1 記載のストレージ装置。

(付記 3)

前記コピー元の記憶装置は、前記指定した記憶装置との通信の成否を判断し、前記パラメータの少なくとも一部を格納したバッファテーブルに判断結果を示す通信成否判断フラグを格納する手段を有する、付記 1 又は 2 記載のストレージ装置。

(付記 4)

前記コピー元の記憶装置は、複数のグループの一括バッファ新規設定処理を行う手段を有する、付記 1 乃至 3 のいずれか 1 項記載のストレージ装置。

(付記 5)

前記コピー元の記憶装置は、前記指定された記憶装置との間の通信が経路の閉塞により失敗した後、経路の復旧した時に一括バッファ新規設定を自動的に継続させる手段を有する、付記 1 乃至 4 のいずれか 1 項記載のストレージ装置。

(付記 6)

前記指定された記憶装置は、前記パラメータに基づいて指定されたバッファサイズのバッファを獲得し、獲得したバッファに関する情報を含むバッファテーブルを作成して前記指定された記憶装置内の記憶部に格納すると共に、前記コピー元の記憶装置へバッファ新規設定処理の正常終了を通知する手段を有する、付記 1 乃至 5 のいずれか 1 項記載のストレージ装置。

(付記 7)

非同期モードでリモートコピーを行う際の書き込み順序性が保証された機能をサポートする分散キャッシュメモリ型のストレージ装置の記憶装置であって、

受領した一括バッファ新規設定コマンドに含まれるパラメータをチェックし、一括バッファ新規設定する記憶装置のグループ及びコピー先の記憶装置を確認して指定された記憶装置と通信を行って前記パラメータを前記指定された記憶装置へ送信する手段と、

前記パラメータに基づく前記指定された記憶装置内のバッファ新規設定処理の結果の通知を受信する手段を有し、

前記パラメータは、前記グループ数、記憶装置を識別する装置識別子、バッファサイズ、及び指定する記憶装置がコピー元であるかコピー先であることを示すバッファ使用用途を含む、ストレージ装置。

【 0 0 5 5 】

以上、本発明を実施例により説明したが、本発明は上記実施例に限定されるものではなく、本発明の範囲内で種々の変形及び改良が可能であることは言うまでもない。

【符号の説明】

【 0 0 5 6 】

- 1 1 制御モジュール
- 1 2 バッファ
- 1 5 CPU
- 2 1 ストレージ装置
- 2 3 - 1 ~ 2 3 - 3 記憶装置
- 2 5 ネットワーク
- 2 6 ストレージ管理 PC
- 3 0 1 ホストサーバ

10

20

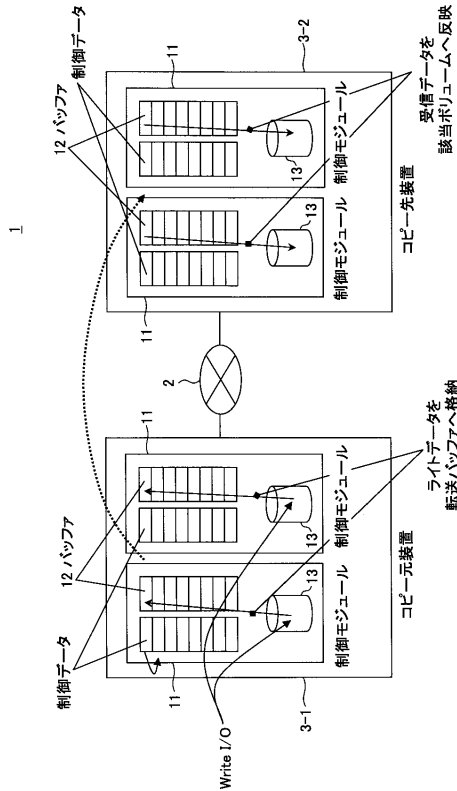
30

40

50

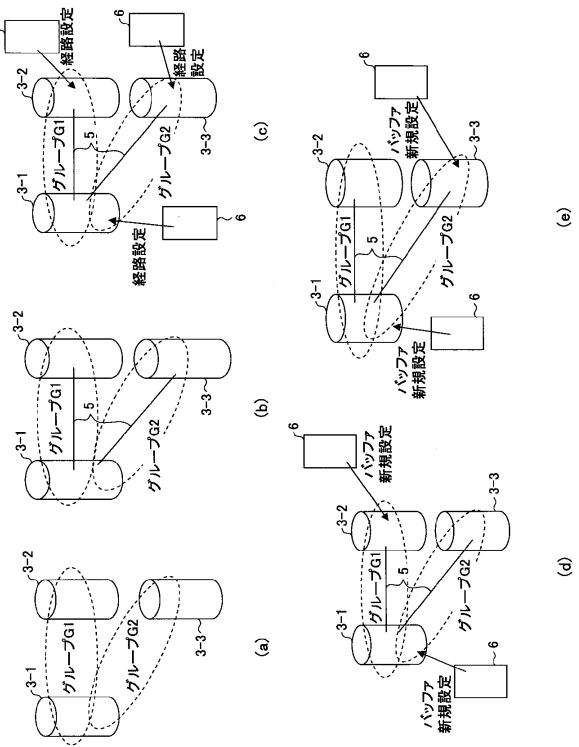
【図1】

従来の分散キャッシュメモリ型のストレージ装置の一例を説明する図



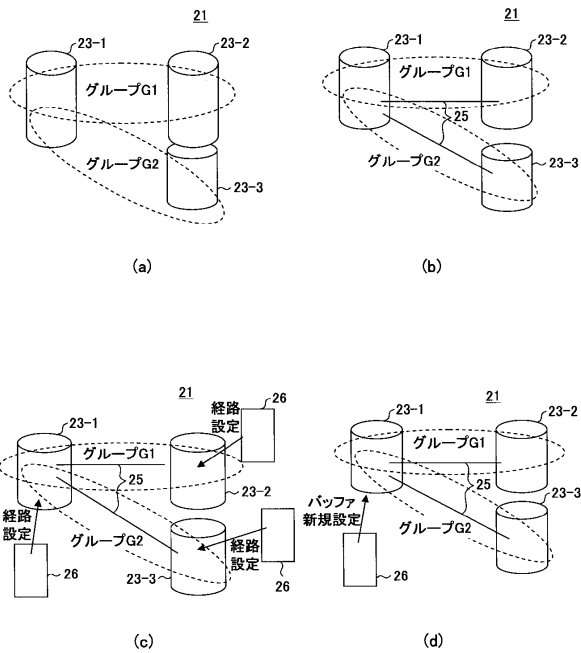
【図2】

RECコンシステンシ・モードで用いられるバッファのバッファ領域新規設定手順の一例を説明する図



【図3】

本発明の一実施例において、分散キャッシュメモリ型のストレージ装置のRECコンシステンシ・モードで用いられるバッファのバッファ領域新規設定手順の一例を説明する図



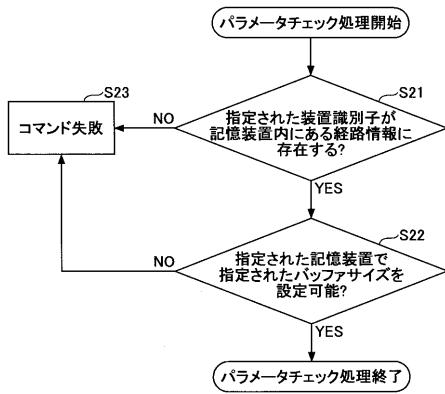
【図4】

REC一括バッファ新規設定コマンドのパラメータの一例を示す図

新規設定グループ数 SGN
相手装置識別子 S11
RECバッファサイズ BS1
RECバッファ使用用途 BU1
:
相手装置識別子 S1n
RECバッファサイズ BSn
RECバッファ使用用途 BUn

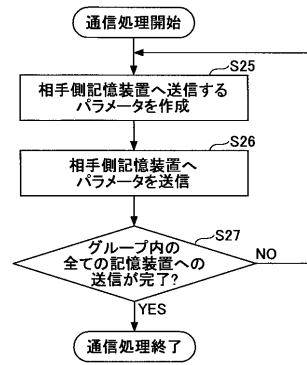
【図5】

パラメータチェック処理を説明するフローチャート



【図6】

通信処理を説明するフローチャート



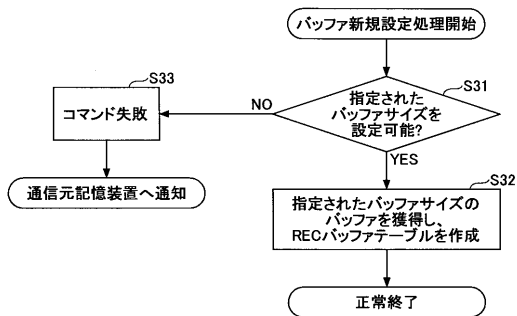
【図7】

ステップS25で作成されるパラメータの一例を示す図

通信元装置識別子 S1s
RECバッファサイズ BS
RECバッファ使用用途 BU

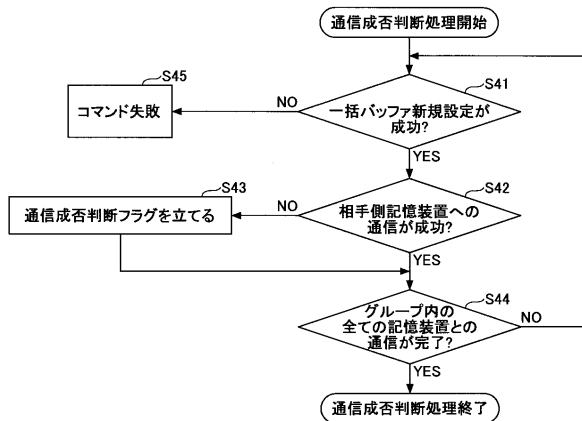
【図8】

処理ST3を説明するフローチャート



【図9】

処理ST4を説明するフローチャート



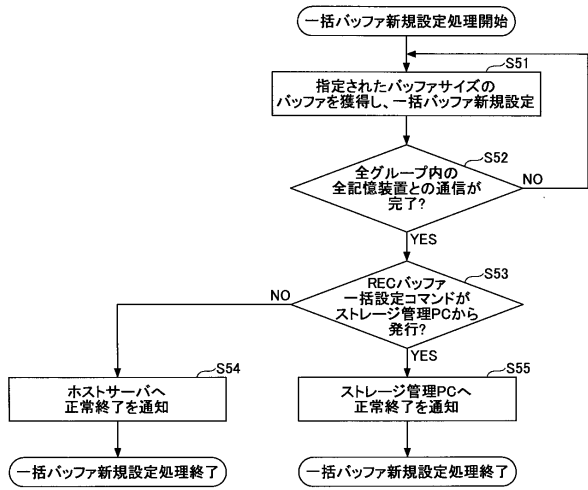
【図10】

RECバッファテーブルの一例を示す図

相手装置識別子 SI1
RECバッファサイズ BS1
RECバッファ使用用途 BU1
通信成否判断フラグ CFJ1
:
相手装置識別子 SIn
RECバッファサイズ BSn
RECバッファ使用用途 BU _n
通信成否判断フラグ CFJ _n

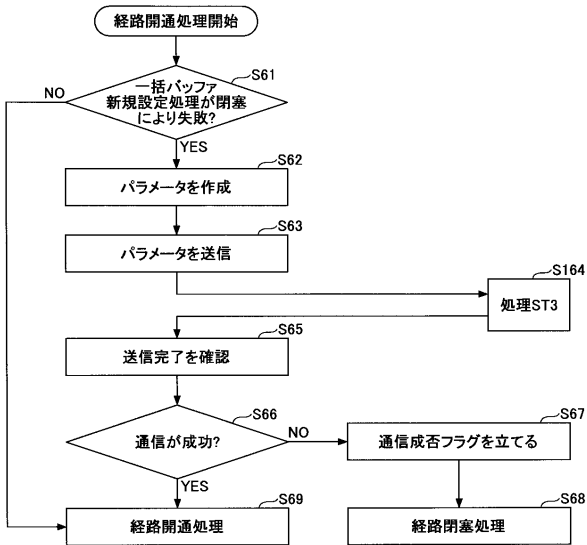
【図11】

処理ST5を説明するフローチャート



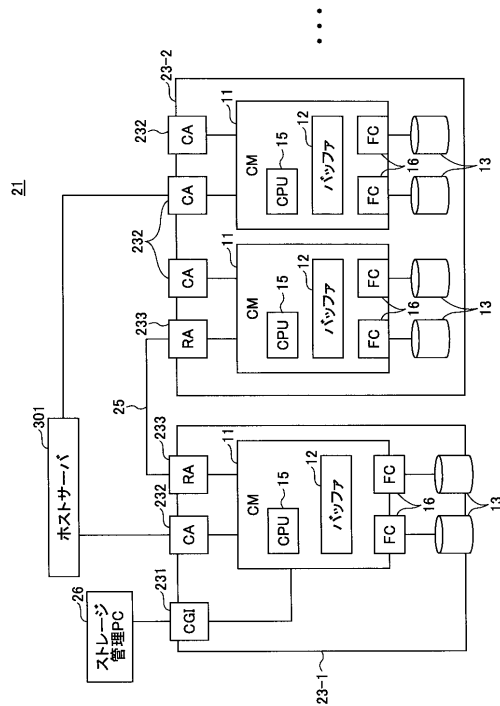
【図12】

処理ST6を説明するフローチャート



【図13】

ストレージ装置のハードウェア構成の一例を示すブロック図



フロントページの続き

- (56)参考文献 特開2005-275537(JP,A)
特表2007-509439(JP,A)
特開2006-260292(JP,A)
特開2003-173295(JP,A)
特開平10-91372(JP,A)
特表2001-518218(JP,A)
特開平7-36651(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06
G06F 13/10