(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2006-309640
(P2006-309640A)
(43)公開日　平成18年11月9日(2006.11.9)

(51)Int.Cl.		F I			テーマコード（参考）
G06F 13/10	(2006.01)	G06F	13/10	340A	5B005
G06F 3/06	(2006.01)	G06F	3/06	302A	5B014
G06F 12/08	(2006.01)	G06F	3/06	304B	5B065
G06F 13/12	(2006.01)	G06F	12/08	501E	
		G06F	12/08	507Z	

審査請求 未請求　請求項の数 20　ＯＬ　（全 25 頁）　最終頁に続く

(21)出願番号　特願2005-133978 (P2005-133978)
(22)出願日　平成17年5月2日 (2005.5.2)

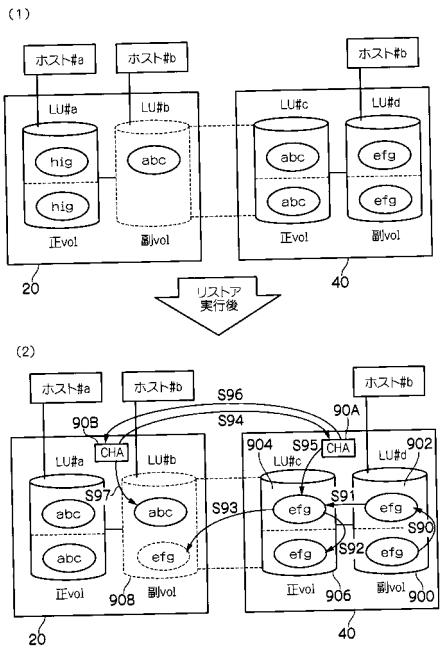
(71)出願人　000005108
株式会社日立製作所
東京都千代田区丸の内一丁目６番６号
(74)代理人　100079108
弁理士　稲葉　良幸
(74)代理人　100093861
弁理士　大賀　眞司
(72)発明者　渡辺　治明
神奈川県小田原市中里３２２番２号　株式会社日立製作所ＲＡＩＤシステム事業部内
(72)発明者　山神　憲司
神奈川県小田原市中里３２２番２号　株式会社日立製作所ＲＡＩＤシステム事業部内

Ｆターム(参考)　5B005 JJ01　MM11　PP21
5B014 EB04　GD13　GD23

最終頁に続く

(54)【発明の名称】記憶制御システム及び記憶制御方法

(57)【要約】
【課題】　複数の記憶制御装置を接続し、一方の記憶制御装置から他方の記憶制御装置にリストアデータを転送する記憶制御システムにおいて、上位装置がリストアデータの転送先の論理ボリュームにアクセスしても、リストアデータに確実にアクセスできるようにした記憶制御システムを提供する。
【解決手段】　第１の記憶制御装置と第２の記憶制御装置とが互いに通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶制御システムであって、前記第１の記憶制御装置は、前記第２の記憶制御装置の論理ボリュームに対応付けられた仮想ボリュームと、当該仮想ボリュームに対応つけられたキャッシュメモリと、当該キャッシュメモリと仮想ボリュームと間のデータ処理を制御する制御部とを備え、前記制御部は、前記論理ボリュームのリストアデータを前記キャッシュメモリに記憶する際に、前記上位装置がアクセスする前記仮想ボリュームに対応する、前記キャッシュメモリの記憶領域をパージする。
【選択図】　図９



【特許請求の範囲】
【請求項１】
第１の記憶制御装置と第２の記憶制御装置とが互いに通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶制御システムであって、
前記第１の記憶制御装置は、
前記第２の記憶制御装置の論理ボリュームに対応付けられた仮想ボリュームと、
当該仮想ボリュームに対応つけられたキャッシュメモリと、
当該キャッシュメモリと仮想ボリュームと間のデータ処理を制御する制御部とを備え、
前記制御部は、前記論理ボリュームのリストアデータを前記キャッシュメモリに記憶する際に、前記上位装置がアクセスする前記仮想ボリュームに対応する、前記キャッシュメモリの記憶領域をパージするように構成された記憶制御システム。
【請求項２】
第１の記憶制御装置と第２の記憶制御装置とが互いに通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶制御システムであって、
前記第１及び第２の記憶制御装置はそれぞれ、
記憶デバイスと、
当該記憶デバイスの論理的な記憶構造であり、前記上位装置からアクセス可能な論理ボリュームと、
前記論理ボリュームのデータを記憶するキャッシュメモリと、
前記記憶デバイス、前記キャッシュメモリ、そして前記論理ボリュームとの間でのデータ処理を制御する制御部と、
を備え、
さらに、前記第１の記憶制御装置は、前記第１の記憶制御装置の記憶デバイスに対応付けされることなく、前記第２の記憶制御装置の論理ボリュームに対応付けられ、かつ前記上位装置がアクセス可能な仮想ボリュームを備え、
前記第１の記憶制御装置の制御部は、前記第２の記憶制御装置の論理ボリュームのリストアデータを前記第１の制御装置のキャッシュメモリに記憶する際に、前記上位装置がアクセスする前記仮想ボリュームの管理領域に対応する当該キャッシュメモリの記憶領域をパージするように構成されてなる、記憶制御システム。
【請求項３】
前記第２の記憶制御装置の論理ボリュームは、コピーペアの関係を形成可能な第１の論理ボリュームと第２の論理ボリュームとを備えてなり、当該第１の論理ボリュームが前記仮想ボリュームに対応付けられている、請求項２記載の記憶制御システム。
【請求項４】
前記第２の記憶制御装置の制御部は、前記リストアの命令を受領すると、前記第２の論理ボリュームの記憶データを前記第１の論理ボリュームにコピーし、さらに、当該第１の論理ボリュームのコピーされたデータを前記仮想ボリュームに転送する、請求項３記載の記憶制御システム。
【請求項５】
前記第１の記憶制御装置の論理ボリュームは前記仮想ボリュームとコピーペアの関係になっている請求項２記載の記憶制御システム。
【請求項６】
前記第１の記憶制御装置の制御部が前記第２の記憶制御装置の制御部に、前記リストアの命令を発行する請求項４記載の記憶制御システム。
【請求項７】
前記上位装置が前記第２の記憶制御装置の制御部に、前記リストア命令を発行する請求項４記載の記憶制御システム。
【請求項８】
前記第１の記憶制御装置の制御部による前記パージが完了後、前記第２の記憶制御装置の制御部が前記リストアを開始する請求項１又は２記載の記憶制御システム。

【請求項９】
前記仮想ボリュームは、前記第１の記憶制御装置のキャッシュメモリの記憶領域との対応関係を制御する情報が設定されたディレクトリ構造を備えており、このディレクトリ構造は、複数の管理領域を備え、各管理領域には、当該管理領域に対応する前記第１の記憶制御装置のキャッシュメモリの記憶領域を特定するための情報が設定されている、請求項１又は２記載の記憶制御システム。
【請求項１０】
前記各管理領域は、前記キャッシュメモリ記憶領域の記憶データの有効又は無効を制御する情報を備えてなる請求項９記載の記憶制御システム。
【請求項１１】
前記第１の記憶制御装置の制御部は、
前記仮想ボリュームの全管理領域の前記特定情報をパージし、
次いで、前記第１の記憶制御装置のキャッシュメモリに前記第２の記憶制御装置の論理ボリュームに保存されているデータを割り当て、
前記上位装置が当該データにアクセスできるように、
構成されてなる請求項１又は２記載の記憶制御システム。
【請求項１２】
前記第１の記憶制御装置の制御部は、
前記仮想ボリュームの一部の管理領域の前記特定情報をパージし、
次いで、前記第１の記憶制御装置のキャッシュメモリに前記第２の記憶制御装置の論理ボリュームに保存されているデータを割り当て、
前記上位装置が当該データにアクセスできるように、
構成されてなる請求項１又は２記載の記憶制御システム。
【請求項１３】
前記第１の記憶制御装置の制御部は、前記仮想ボリュームに全体世代番号を、当該仮想ボリュームの各管理領域に個別の世代番号を設定し、前記上位装置がアクセスした前記仮想ボリュームの管理領域の世代番号と前記全体番号を比較し、両者が一致している場合には、当該管理領域に対応するキャッシュメモリ記憶領域のデータを前記リストアデータと判定し、両者が一致していない場合には、前記管理領域に対応するキャッシュメモリの記憶領域のデータをパージし、次いで、当該記憶領域に前記リストアデータを記憶するようにした請求項請求項１２記載の記憶制御システム。
【請求項１４】
前記第１の記憶制御装置の制御部は、前記仮想ボリュームの代表世代番号を、前記上位装置からの前記パージのコマンドを受ける都度、更新するようにした請求項１３記載の記憶制御システム。
【請求項１５】
前記第２記憶制御装置の論理ボリュームに他の上位装置が接続され、当該上位装置からのリストアデータが当該論理ボリュームに記憶されている請求項２記載の記憶制御システム。
【請求項１６】
前記仮想ボリュームとコピーペアの関係を持つ他の論理ボリュームが前記第１の記憶制御装置に備わっている、請求項３記載の記憶制御システム。
【請求項１７】
前記パージ処理は、前記第１の記憶制御装置の制御部が、前記仮想ボリュームの管理領域に前記キャッシュメモリの記憶領域との対応関係が無いことを示す制御コードを設定するものである請求項９記載の記憶制御システム。
【請求項１８】
前記パージ処理は、前記第１の記憶制御装置の制御部が、前記管理領域に、前記キャッシュメモリ記憶領域の記憶データが無効であることを示す制御コードを設定するものである請求項２記載の記憶制御システム。

【請求項１９】
前記第１の記憶制御装置の制御部は、前記パージ処理された、前記キャッシュメモリの記憶領域に、前記リストアデータを記憶する請求項１又は２記載の記憶制御システム。
【請求項２０】
第１の記憶制御装置と第２の記憶制御装置とが互いに通信可能に接続された記憶制御システムに対して適用され、上位装置からのリクエストに応じたデータ処理を行う記憶制御方法であって、
前記第１の記憶制御装置は、
前記第２の記憶制御装置の論理ボリュームに対応付けられた仮想ボリュームと、
当該仮想ボリュームに対応つけられたキャッシュメモリと、
当該キャッシュメモリと仮想ボリュームと間のデータ処理を制御する制御部とを備え、
前記制御部は、前記論理ボリュームのリストアデータを前記キャッシュメモリに記憶する際に、前記上位装置がアクセスする前記仮想論理ボリュームに対応する前記キャッシュメモリの記憶領域をパージするステップを実行する記憶制御方法。

【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、ホストコンピュータが接続されて、当該ホストコンピュータとの間で記憶データの通信を行う記憶制御装置を備えてなる記憶制御システムに関するものである。
【背景技術】
【０００２】
例えば、データセンタ等のような大規模なデータを取り扱うデータベースシステムでは、ホストコンピュータとは別に構成された記憶制御システムを用いてデータを管理する。この記憶制御システムは、例えば、ディスクアレイ装置を備えて構成される。ディスクアレイ装置は、多数の記憶デバイスをアレイ状に配設して構成されるもので、例えば、RAID（Redundant Array of Independent Inexpensive Disks）に基づいて構築されている。記憶デバイス群が提供する物理的な記憶領域上には少なくとも１つ以上の論理的な記憶領域である論理ボリュームが形成され、この論理ボリュームが、上位装置としてのホストコンピュータ（より詳しくは、ホストコンピュータ上で稼働するデータベースなどのプログラム）に提供される。ホストコンピュータは、所定のコマンドをディスクアレイ装置に送信することにより、論理ボリュームに対してデータの書込み、読み出しができる。
【０００３】
この種の記憶制御システムにおいて、処理の実行中に発生した障害などによって破壊されたデータをバックアップデータから高速に回復することが行われる。例えば、特開２００１－２１６１８５号に記載されたディスクアレイ装置では、データを復旧させるために次のように構成されている。ホストは、ペア形成要求およびペア分割要求をディスクアレイ装置に発行し、ボリュームのスナップショットを作成した後、処理を開始する。ディスクサブシステムは、スナップショットの作成以降、更新が行われたデータの位置を差分情報として記憶する。ホストからのデータ復旧要求を受領すると、ディスクアレイ装置は、差分情報に基づいて、更新が行われた箇所のデータをスナップショットからリストアする。
【特許文献１】特開２００１－２１６１８５
【発明の開示】
【発明が解決しようとする課題】
【０００４】
前記従来のディスクサブシステムは、一台の記憶制御装置内でのリストアに関するものであり、複数の記憶制御装置を接続し、一方の記憶制御装置から他方の記憶制御装置にデータをリストアする際のデータ復旧のことまでは配慮していない。複数の記憶制御装置間でのデータのリストアは、ホストのパフォーマンスを落とさないようにするために、ホス

トを経由することなく記憶制御装置間で行われることが望まれるが、リストアデータの送達を受けた記憶制御装置のキャッシュメモリ上にリストア前のデータが残っていると、リストア前のデータがキャッシュヒットしてホストに転送されてしまう。そのため、ホストが、他の記憶制御装置の記憶資源に割り当てられた論理ボリュームにアクセスすると、リストアされるべきデータではなく、リストア前のデータがホストに見えてしまう問題がある。

【０００５】

そこで、本発明は、複数の記憶制御装置を接続し、一方の記憶制御装置から他方の記憶制御装置にリストアデータを転送する記憶制御システムにおいて、上位装置がリストアデータの転送先の論理ボリュームにアクセスしても、リストアデータに確実にアクセスできるようにした記憶制御システム及び記憶制御方法を提供することを目的とする。本発明の他の目的は、キャッシュメモリのパージに要する時間を短縮し、ホストが早くリストアデータにアクセスできるようにした記憶制御システムを提供することにある。

【課題を解決するための手段】

【０００６】

前記目的を達成するために、本発明は、記憶制御装置のキャッシュメモリに残っているリストア前のデータをパージした後、キャッシュメモリの記憶領域にリストアデータを記憶させるように構成されている。

【０００７】

すなわち、本発明は、第１の記憶制御装置と第２の記憶制御装置とが互いに通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶制御システムであって、前記第１の記憶制御装置は、前記第２の記憶制御装置の論理ボリュームに対応付けられた仮想ボリュームと、当該仮想ボリュームに対応つけられたキャッシュメモリと、当該キャッシュメモリと仮想ボリュームと間のデータ処理を制御する制御部とを備え、前記制御部は、前記論理ボリュームのリストアデータを前記キャッシュメモリに記憶する際に、前記上位装置がアクセスする前記仮想論理ボリュームに対応する、前記キャッシュメモリの記憶領域をパージするように構成されたことを特徴とするものである。

【０００８】

さらに本発明は、第１の記憶制御装置と第２の記憶制御装置とが互いに通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶制御システムであって、前記第１及び第２の記憶制御装置はそれぞれ、記憶デバイスと、当該記憶デバイスの論理的な記憶構造であり、前記上位装置からアクセス可能な論理ボリュームと、前記論理ボリュームのデータを記憶するキャッシュメモリと、前記記憶デバイス、前記キャッシュメモリ、そして前記論理ボリュームとの間でのデータ処理を制御する制御部と、を備え、さらに、前記第１の記憶制御装置は、前記第１の記憶制御装置の記憶デバイスに対応付けされることなく、前記第２の記憶制御装置の論理ボリュームに対応付けられ、かつ前記上位装置がアクセス可能な仮想ボリュームを備え、前記第１の記憶制御装置の制御部は、前記第２の記憶制御装置の論理ボリュームのリストアデータを前記第１の制御装置のキャッシュメモリに記憶する際に、前記上位装置がアクセスする前記仮想ボリュームの管理領域に対応する当該キャッシュメモリの記憶領域をパージするように構成されてなる、ことを特徴とするものである。

【０００９】

本発明の好適な形態では、前記第２の記憶制御装置の論理ボリュームは、コピーペアの関係を形成可能な第１の論理ボリュームと第２の論理ボリュームとを備えてなり、当該第１の論理ボリュームが前記仮想ボリュームに対応付けられている。第２の論理ボリュームから第１の論理ボリュームへデータのリストアが実施される。

【００１０】

前記第２の記憶制御装置の制御部は、前記リストアの命令を受領すると、前記第２の論理ボリュームの記憶データを前記第１の論理ボリュームにコピーし、さらに、当該第１の論理ボリュームのコピーされたデータを前記仮想ボリュームに転送する。前記第１の記憶

制御装置の論理ボリュームは前記仮想ボリュームとコピーペアの関係になっている。前記第１の記憶制御装置の制御部が前記第２の記憶制御装置の制御部に、前記リストアの命令を発行する。前記上位装置が前記第２の記憶制御装置の制御部に、前記リストア命令を発行する。

【００１１】

前記第１の記憶制御装置の制御部による前記パージが完了後、前記第２の記憶制御装置の制御部が前記リストアを開始する。前記仮想ボリュームは、前記第１の記憶制御装置のキャッシュメモリの記憶領域との対応関係を制御する情報が設定されたディレクトリ構造を備えており、このディレクトリ構造は、複数の管理領域を備え、各管理領域には、当該管理領域に対応する前記第１の記憶制御装置のキャッシュメモリの記憶領域を特定するための情報が設定されている。前記各管理領域は、前記キャッシュメモリ記憶領域の記憶データの有効又は無効を制御する情報を備えてなる。前記第１の記憶制御装置の制御部は、前記仮想ボリュームの全管理領域の前記特定情報をパージし、次いで、前記第１の記憶制御装置のキャッシュメモリに前記第２の記憶制御装置の論理ボリュームに保存されているデータを割り当て、前記上位装置が当該データにアクセスできるように、構成されてなる。前記第１の記憶制御装置の制御部は、前記仮想ボリュームの一部の管理領域の前記特定情報をパージし、次いで、前記第１の記憶制御装置のキャッシュメモリに前記第２の記憶制御装置の論理ボリュームに保存されているデータを割り当て、前記上位装置が当該データにアクセスできるように、構成されてなる。

【００１２】

前記第１の記憶制御装置の制御部は、前記仮想ボリュームに全体世代番号を、当該仮想ボリュームの各管理領域に個別の世代番号を設定し、前記上位装置がアクセスした前記仮想ボリュームの管理領域の世代番号と前記全体番号を比較し、両者が一致している場合には、当該管理領域に対応するキャッシュメモリ記憶領域のデータを前記リストアデータと判定し、両者が一致していない場合には、前記管理領域に対応するキャッシュメモリの記憶領域のデータをパージし、次いで、当該記憶領域に前記リストアデータを記憶するようにした。前記第１の記憶制御装置の制御部は、前記仮想ボリュームの代表世代番号を、前記上位装置からの前記パージのコマンドを受ける都度、更新するようにした。前記第２記憶制御装置の論理ボリュームに他の上位装置が接続され、当該上位装置からのリストアデータが当該論理ボリュームに記憶されている。前記仮想ボリュームとコピーペアの関係を持つ他の論理ボリュームが前記第１の記憶制御装置に備わっている。

【００１３】

前記パージ処理は、前記第１の記憶制御装置の制御部が、前記仮想ボリュームの管理領域に前記キャッシュメモリの記憶領域との対応関係が無いことを示す制御コードを設定するものである。前記パージ処理は、前記第１の記憶制御装置の制御部が、前記管理領域に、前記キャッシュメモリ記憶領域の記憶データが無効であることを示す制御コードを設定するものである。前記第１の記憶制御装置の制御部は、前記パージ処理された、前記キャッシュメモリの記憶領域に、前記リストアデータを記憶する。

【発明の効果】

【００１４】

以上説明したように、本発明によれば、複数の記憶制御装置を接続し、一方の記憶制御装置から他方の記憶制御装置にリストアデータを転送する記憶制御システムにおいて、上位装置がリストアデータの転送先の論理ボリュームにアクセスしても、リストアデータに確実にアクセスできるようになった。

【発明を実施するための最良の形態】

【００１５】

本発明の実施形態を以下に説明する。以下に説明する代表的な記憶制御システムは、第１の記憶制御装置に仮想的に論理記憶領域を設定し、この仮想領域に第１の記憶制御装置の外部に存在する第２の記憶制御装置の論理記憶領域を対応付けて(マッピングして)、第１の記憶制御装置が第２の記憶制御装置の記憶領域をあたかも自分の記憶領域であるかの

ようにしてこれを上位装置に提供している。

【００１６】

図１は、記憶制御システムの要部の構成を示すブロック図である。上位装置としてのホスト装置１０は、例えば、ＣＰＵ（Central Processing Unit）やメモリ等の情報処理資源を備えたコンピュータ装置であり、例えば、パーソナルコンピュータ、ワークステーション、メインフレーム等として構成される。ホスト装置１０は、例えば、キーボード、スイッチやポインティングデバイス、マイクロフォン等の情報入力装置（図示せず）と、例えば、モニタディスプレイやスピーカー等の情報出力装置（図示せず）とを備えている。さらに、ホスト装置１０には、例えば、第１の記憶制御装置２０が提供する記憶領域を使用するデータベースソフトウェア等のアプリケーションプログラム１１と、通信ネットワークＣＮ１を介して第１の記憶制御装置２０にアクセスするためのアダプタ１２とが設けられている。

【００１７】

ホスト装置１０は、スイッチ回路ＳＷを備えて構成される通信ネットワークＣＮ１を介して第１の記憶制御装置２０に接続されている。通信ネットワークＣＮ１としては、例えば、ＬＡＮ、ＳＡＮ、インターネット、専用回線、公衆回線等を場合に応じて適宜用いることができる。ＬＡＮを介するデータ通信は、例えば、TCP/IP（Transmission Control Protocol/Internet Protocol）プロトコルに従って行われる。ホスト装置１０がＬＡＮを介して第１の記憶制御装置２０に接続される場合、ホスト装置１０は、ファイル名を指定してファイル単位でのデータ入出力を要求する。一方、ホスト装置１０がＳＡＮを介して第１の記憶制御装置２０等に接続される場合、ホスト装置１０は、ファイバチャネルプロトコルに従って、複数のディスク記憶装置（ディスクドライブ）により提供される記憶領域のデータ管理単位であるブロックを単位としてデータ入出力を要求する。通信ネットワークＣＮ１がＬＡＮである場合、アダプタ１２は、例えばＬＡＮ対応のネットワークカードである。通信ネットワークＣＮ１がＳＡＮの場合、アダプタ１２は、例えばホストバスアダプタである。

【００１８】

スイッチ回路ＳＷは通信ネットワークに接続されたルータや交換機を備えて構成されている。スイッチ回路ＳＷは、第１の記憶制御装置２０のエクスターナルポート２１Ａに、第２の記憶制御装置４０のターゲットポート４１及び第１の記憶制御装置のターゲットポート２１Ｂを切り換えて接続できるように構成されている。なお、第１の記憶制御装置２０は、エクスターナルポート２１Ａとスイッチ回路ＳＷを介して自筐体のターゲットポート２１Ｂに接続できるように構成されているために、所謂自己ループ式に分類されるものである。各ポート及びスイッチ回路にはネットワーク上のアドレスが設定されている。

【００１９】

第１の記憶制御装置２０は、例えば、ディスクアレイサブシステムとして構成されるものである。但し、これに限らず、第１の記憶制御装置２０を、高機能化されたインテリジェント型のファイバチャネルスイッチとして構成することもできる。第１の記憶制御装置２０は、後述のように、第２の記憶制御装置４０の有する記憶資源を自己の論理ボリューム（Logical Unit）としてホスト装置１０に提供する。

【００２０】

第１の記憶制御装置２０は、コントローラ部と記憶装置部とに大別することができ、コントローラ部は、例えば、複数のチャネルアダプタ（ＣＨＡ）２１と、複数のディスクアダプタ（ＤＫＡ）２２と、サービスプロセッサ（ＳＶＰ）２３と、キャッシュメモリ２４と、共有メモリ２５と、接続部２６とを備えている。このコントローラ部が請求項記載の制御部に対応する。

【００２１】

チャネルアダプタ（ＣＨＡ）２１は、ホスト装置１０との間のデータ通信を行うものである。他のチャネルアダプタ２１は、第１の記憶制御装置のエクスターナルポート２１Ａ、第２の記憶制御装置のターゲットポート４１を介して、第２の記憶制御装置４０の内部

論理ボリュームとデータ通信を行うものである。各チャネルアダプタ２１は、それぞれマイクロプロセッサやメモリ等を備えたマイクロコンピュータシステムとして構成されており、ホスト装置１０から受信した各種コマンドを解釈して実行する。各チャネルアダプタ２１には、それぞれを識別するためのネットワークアドレス（例えば、ＩＰアドレスやＷＷＮ）が割り当てられているので、各チャネルアダプタ２１は、それぞれが個別にＮＡＳ（Network Attached Storage）として振る舞うことができるようになっている。複数のホスト装置１０が存在する場合、各チャネルアダプタ２１はホスト装置毎に設けられ、各ホスト装置１０からの要求をそれぞれ個別に受け付けることができるように構成されている。

【００２２】

各ディスクアダプタ（ＤＫＡ）２２は、記憶装置３０の記憶デバイス３１，３２との間のデータ授受を行うものである。各ディスクアダプタ２２は、記憶デバイス３１，３２に接続するための通信ポート２２Ａを備えている。また、各ディスクアダプタ２２は、マイクロプロセッサやメモリ等を備えたマイクロコンピュータシステムとして構成されている。各ディスクアダプタ２２は、チャネルアダプタ２１がホスト装置１０から受信したデータを、ホスト装置１０からのリクエストに基づいて（書込み命令）、所定の記憶デバイス３１，３２の所定のアドレスに書込み、また、ホスト装置１０からのリクエストに基づいて（読み出し命令）、所定の記憶デバイス３１，３２の所定のアドレスからデータを読み出し、ホスト装置１０に送信させる。記憶デバイス３１，３２との間でデータ入出力を行う場合、各ディスクアダプタ２２は、論理的なアドレスを物理的なアドレスに変換する。各ディスクアダプタ２２は、記憶デバイス３１，３２がRAIDに従って管理されている場合は、RAID構成に応じたデータアクセスを行う。

【００２３】

サービスプロセッサ（ＳＶＰ）２３は、装置全体の作動を制御するものである。ＳＶＰ２３には、例えば、管理用クライアント（図示せず）が接続される。ＳＶＰ２３は、装置内の障害発生を監視して管理用クライアントに表示させたり、管理用クライアントからの指令に基づいて記憶ディスクの閉塞処理等を指示するようになっている。さらに、ＳＶＰ２３の管理用クライアントは後述のように仮想ボリュームを定義する処理を実行する。管理用クライアントは、例えばＪＡＶＡ(登録商標）アプレット上に管理用プログラムが実装して構成されている。

【００２４】

キャッシュメモリ２４は、ホスト装置１０から受信したデータや、記憶デバイス３１，３２から読み出したデータを一時的に記憶するものである。共有メモリ２５には、第1の記憶制御装置の作動に使用するための制御情報等が格納される。また、共有メモリ２５には、ワーク領域が設定されるほか、後述するマッピングテーブルＴｍ等の各種テーブル類も格納される。また、キャッシュメモリ１３０と共有メモリ１４０とは、それぞれ別々のメモリとして構成することもできるし、同一のメモリの一部の記憶領域をキャッシュ領域として使用し、他の記憶領域を制御領域として使用することもできる。なお、記憶デバイス３１，３２のいずれか１つあるいは複数を、キャッシュ用のディスクとして使用してもよい。

【００２５】

接続部２６は、各チャネルアダプタ２１，各ディスクアダプタ２２，ＳＶＰ２３，キャッシュメモリ２４，共有メモリ２５を相互に接続させる。接続部２６は、例えば、高速スイッチング動作によってデータ伝送を行う超高速クロスバスイッチ等のような高速バスとして構成することができる。

【００２６】

記憶装置３０は、複数の記憶デバイス３１を備えている。記憶デバイス３１としては、例えば、ハードディスク、フレキシブルディスク、磁気テープ、半導体メモリ、光ディスク等のようなデバイスを用いることができる。また、例えば、FC（Fibre Channel）ディスクやSATA（Serial AT Attachment）ディスク等のように、異種類のディスクを記憶装置

３０内に混在させることもできる。記憶装置３０内に点線で示される記憶デバイス３２は、第２の記憶制御装置４０の有する記憶デバイス４２を第１の記憶制御装置２０側に取り込んだ状態を示すものである。即ち、第１の記憶制御装置２０から見て外部に存在する記憶デバイス４２を、第１の記憶制御装置２０の内部記憶デバイスかのように認識できるようにし、ホスト装置１０に外部記憶デバイス４２の記憶資源を提供する。このことは後述のように、第1の記憶制御装置２０内の中間論理記憶領域である仮想ボリュームに第２の記憶制御装置４０の論理ボリュームをマッピングすることにより可能になる。キャッシュメモリ２４の記憶空間を利用して、仮想ボリュームが構築される。第1の記憶制御装置２０に形成された仮想ボリュームは、第1の記憶制御装置２０内の実論理ボリュームとともにホスト装置１０に認識される。なお、図１に点線で示される外部の記憶デバイス３２とのデータ交換は第２の記憶制御装置４０のディスクアダプタによって実行される。

【００２７】

第２の記憶制御装置４０は、通信ポート（ターゲットポート）４１と記憶デバイス４２とを備えている。このほか、チャネルアダプタやディスクアダプタ等を備えることもできる。第２の記憶制御装置４０は、スイッチ回路ＳＷを介して第１の記憶制御装置２０に接続されており、第２の記憶制御装置４０の記憶デバイス４２は、第１の記憶制御装置２０の内部記憶デバイスとして扱われるようになっている。スイッチ回路ＳＷには複数の外部記憶制御装置４０Ａを接続可能である。

【００２８】

続いて図２について説明する。図２は、第１の記憶制御装置２０及び第２の記憶制御装置４０の１つの記憶構造の概略を示すブロック図である。まず、第１の記憶制御装置の構成から先に説明する。

【００２９】

第１の記憶制御装置の記憶構造は、例えば、物理的記憶階層と論理的記憶階層とに大別することができる。物理的記憶階層は、物理的なディスクであるPDEV（Physical Device）１６１により構成される。PDEVは、ディスクドライブに該当する。

【００３０】

論理的記憶階層は、複数の（例えば２種類の）階層から構成することができる。一つの論理的階層は、VDEV（Virtual Device）１６２と、VDEV１６２のように扱われる仮想的なVDEV（以下、「Ｖ－VOL」とも呼ぶ）１６３とから構成可能である。他の一つの論理的階層は、LDEV（Logical Device）１６４から構成することができる。

【００３１】

VDEV１６２は、例えば、４個１組（３Ｄ＋１Ｐ）、８個１組（７Ｄ＋１Ｐ）等のような所定数のPDEV１６１をグループ化して構成される。グループに属する各PDEV１６１がそれぞれ提供する記憶領域が集合して一つのRAID記憶領域が形成されている。このRAID記憶領域がVDEV１６２となる。

【００３２】

VDEV１６２が物理的な記憶領域上に構築されるのと対照的に、Ｖ－VOL１６３は、物理的な記憶領域を必要としない仮想的な中間記憶デバイスである。Ｖ－VOL１６３は、物理的な記憶領域に直接関係づけられるものではなく、第２の記憶制御装置４１のLU（Logical Unit）２５０をマッピングするための受け皿となる仮想的な存在である。

【００３３】

LDEV１６４は、VDEV１６２またはＶ－VOL１６３上に、それぞれ少なくとも一つ以上設けることができる。LDEV１６４は、例えば、VDEV１６２を固定長で分割することにより構成することができる。ホスト１０がオープン系ホストの場合、LDEV１６４がLU１６５にマッピングされることにより、ホスト１０は、LDEV１６４を一つの物理的なディスクとして認識する。オープン系のホストは、LUN（Logical Unit Number）や論理ブロックアドレスを指定することにより、所望のLDEV１６４にアクセスする。なお、メインフレーム系ホストの場合は、LDEV１６４を直接認識する。

【００３４】

LU１６５は、SCSIの論理ユニットとして認識可能なデバイスである。各LU１６５は、ターゲットポート２１Ｂを介してホスト１０に接続される。各LU１６５には、少なくとも一つ以上のLDEV１６４をそれぞれ関連付けることができる。一つのLU１６５に複数のLDEV１６４を関連付けることにより、LUサイズを仮想的に拡張することもできる。

【００３５】

CMD（Command Device）１６６は、ホスト１０上で稼働するI/O制御プログラムと記憶制御装置２０のコントローラ（CHA２１，DKA２２）との間で(図１参照)、コマンドやステータスを受け渡すために使用される専用のLUである。ホスト１０からのコマンドは、CMD１６６に書き込まれる。記憶制御装置２０のコントローラは、CMD１６６に書き込まれたコマンドに応じた処理を実行し、その実行結果をステータスとしてCMD１６６に書き込む。ホスト装置１０は、CMD１６６に書き込まれたステータスを読出して確認し、次に実行すべき処理内容をCMD１６６に書き込む。このようにして、ホスト装置１０は、CMD１６６を介して、記憶制御装置２０に各種の指示を与えることができる。

【００３６】

なお、ホスト装置１０から受信したコマンドを、CMD１６６に格納することなく、処理することもできる。また、実体のデバイス（LU）を定義せずに、CMDを仮想的なデバイスとして生成し、ホスト装置１０からのコマンドを受け付けて処理するように構成してもよい。即ち、例えば、CHA２１は、ホスト装置１０から受信したコマンドを共有メモリ２５に書き込み、この共有メモリ２５に記憶されたコマンドを、CHA２１又はDKA２２が処理する。その処理結果は共有メモリ２５に書き込まれ、CHA２１からホスト装置１０に送信される。

【００３７】

さて、第１の記憶制御装置２０の外部接続用のエクスターナルポート（External Port）２１Ａには、既述の通りスイッチ回路ＳＷを介して第２ストレージ装置４０のターゲットポート４１又は第1の記憶制御装置２０のターゲットポート２１Ｂが接続可能である。

【００３８】

第２の記憶制御装置４０は、複数のPDEV２２０と、PDEV２２０の提供する記憶領域上に設定されたVDEV２３０と、VDEV２３０上に少なくとも一つ以上設定可能なLDEV２４０とを備えている。各LDEV２４０は、LU２５０にそれぞれ関連付けられている。

【００３９】

そして、本実施形態では、第２の記憶制御装置４０のLU２５０（即ち、LDEV２４０）は、仮想的な中間記憶デバイスであるV-VOL１６３にマッピングされており、第１の記憶制御装置２０の内部論理ボリュームとして取り扱われる。

【００４０】

例えば、第２の記憶制御装置４０の「LDEV１」，「LDEV２」は、「LU１」，「LU２」を介して、第１の記憶制御装置２０の「V－VOL１」，「V－VOL２」にそれぞれマッピングされている。そして、「V－VOL１」，「V－VOL２」は、それぞれ「LDEV３」，「LDEV４」にマッピングされ、「LU３」，「LU４」を介して、ホスト装置１０が第２の記憶制御装置の記憶領域にアクセス可能になっている。

【００４１】

なお、VDEV１６２，V-VOL１６３は、RAID構成を適用することができる。即ち、１つのディスクドライブ１６１を複数のVDEV１６２，V-VOL１６３に割り当てることもできるし（スライシング）、複数のディスクドライブ１６１から１つのVDEV１６２，V-VOL１６３を形成することもできる（ストライピング）。

【００４２】

そして、第１の記憶制御装置２０の「LDEV１」または「LDEV２」が第１の記憶制御装置２０に記憶領域を持つ、実論理ボリューム１９０に該当する。第１の記憶制御装置２０の「LDEV３」または「LDEV４」が、第１の記憶制御装置に記憶領域を持たず、第２の記憶制御装置に記憶領域を持つ仮想ボリューム１９１に該当する。第２の記憶制御装置４０の「LDEV１」または「LDEV２」が、仮想ボリュームに対応する（仮想ボリュームがマッピング

された）論理ボリューム２６０に該当する。このように、実論理ボリューム１９０は、第１の記憶制御装置１内に設けられた物理的な記憶デバイス（例えば、ディスクドライブ等）に基づいて形成されるものである。仮想ボリューム１９１は、仮想的な存在であり、データを記憶する実体は、第２の記憶制御装置４０内に存在する。即ち、第１の記憶制御装置２０が有する記憶階層の所定の層に、第２の記憶制御装置４０の有する論理ボリューム２６０がマッピングされるように、仮想ボリューム１９１が構築されている。

【００４３】

図２によれば、第１の記憶制御装置２０のＣＨＡ２１は、エクスターナルポート２１Ａから参照できるボリュームを外部デバイスとして認識する。この外部記憶デバイスにマッピングされる既述の仮想ボリュームが第１の記憶制御装置２０内に定義される。この定義動作は、既述のＳＶＰ２３の管理クライアントである、例えばJAVA（登録商標）アプレットによって実行される。この定義情報はマッピングテーブルＴｍとして共有メモリ２５上に置かれる。但し、図２の符号３００に示すように、第１の記憶制御装置の論理ボリューム１９０に仮想ボリュームをマッピングすることはできない。

【００４４】

次に仮想ボリュームを第１の記憶制御装置２０に設定するための動作について説明する。図３はこの設定動作のフローを示す。管理クライアント上のJAVA（登録商標）アプレットは、ＳＶＰにエクスターナルポート２１Ａ（図１）から参照できる論理ボリューム情報取得(ディスカバリー)をリクエストする（Ｓ１）。このリクエストを受けたＳＶＰは、第１の記憶制御装置のCHA２１（図１）に、ディスカバリーの実行を指令する（Ｓ２）。CHAは共有メモリ２５を参照してスイッチ回路ＳＷのＩＰアドレスを入手してエクスターナルポート２１Ａを介してスイッチ回路ＳＷにアクセスし、このスイッチ回路から参照できる全てのボリュームの情報を取得する。すなわち、CHAは自己の筐体の内部論理ボリュームか他の記憶制御装置の内部論理ボリュームかを区別することなく、エクスターナルポート２１Ａに接続可能な内部論理ボリュームのディスカバリーを実行する（Ｓ３）。図１に示すように、CHA２１がエクスターナルポート２１Ａに接続されるスイッチ回路ＳＷを介して参照できるボリュームは、第１の記憶制御装置２０の論理ボリュームと第２の記憶制御装置４０の論理ボリュームである。

【００４５】

次いで、第１の記憶制御装置２０のＣＨＡ２１及び第２の記憶制御装置４０のＣＨＡ（図示せず）は、それぞれの共有メモリを参照して、それぞれの内部論理ボリュームの特性データを取得する（Ｓ４）。この特性データを受けたＣＨＡ２１はＳＶＰ３２を介して管理クライアントのJAVA（登録商標）アプレットに特性データを送信する（Ｓ５、Ｓ６）。JAVA（登録商標）アプレット上には、特性データ中の「ストレージのベンダー名、製品型名、製造番号」等の情報からディスカバリーによって参照された複数の内部論理ボリュームのそれぞれについて、抽出された内部論理ボリュームが属する筐体が、ＣＨＡ２１が属する自己の記憶制御装置か或いは外部記憶制御装置かを判断するストレージ管理用アプリケーションが実装されている。

【００４６】

この管理用アプリケーションは、ディスカバリーによって抽出された複数の内部論理ボリュームについて、既述の仮想ボリュームに対する候補として良いか否かのフィルタリングを行うモジュールを備えている。管理用アプリケーションは、このフィルタリングモジュールによって複数の論理ボリュームについての各抽出情報をフィルタリング（Ｓ７）し、合致した条件の論理ボリュームについて仮想ボリュームのマッピング先の候補群として画面表示処理が行われる。

【００４７】

本実施形態では、既述のとおり、第１の記憶制御装置２０の内部論理ボリュームが、仮想ボリュームがマッピングされるボリュームとしての候補群として表示されないように管理用アプリケーションがプログラムされている。次いで、管理用クライアントはＳＶＰに仮想ボリュームの定義処理の実行をリクエストする（Ｓ８、Ｓ９）。ＳＶＰはＣＨＡに仮

想ボリュームの定義処理、すなわち、既述のマッピングテーブル作成の処理動作を実行させ（Ｓ１０）、その処理結果を受けてこれを管理用クライアントに送る（Ｓ１１，Ｓ１２）。管理用クライアントのユーザは自筐体（第１の記憶制御装置２０）の内部論理ボリュームが自筐体内の仮想ボリュームのマッピング先として表示されないので、自筐体の内部論理ボリュームに対して仮想ボリュームを設定することができない。このように、ＳＶＰ及び管理用クライアントによって仮想ボリューム設定制御が実現される。この設定制御処理はＣＨＡ或いはＤＫＡのプロセッサーによっても実現可能である。

【００４８】

次に、仮想ボリュームの定義処理について詳細を説明する。図４は、既述のマッピングテーブルＴｍの構造の一例である。マッピングテーブルＴｍは、例えば、ＶＤＥＶをそれぞれ識別するためのＶＤＥＶ番号と外部の記憶デバイスの情報とをそれぞれ対応付けることにより、構成ができる。外部デバイス情報としては、例えば、デバイス識別情報と、記憶デバイスの記憶容量と、デバイスの種別を示す情報と（例えば、テープ系デバイスかディスク系デバイスか等）、記憶デバイスへのパス情報とを含んで構成することができる。また、パス情報は、各通信ポート(図１の２１Ａ,２１Ｂ)に固有の識別情報（ＷＷＮ）と、図２のＬＵを識別するためのＬＵ番号とを含んで構成できる。

【００４９】

なお、図４中に示すデバイス識別情報やＷＷＮ等は、説明の便宜上の値であって特に意味はない。また、図４中の下側に示すＶＤＥＶ番号「３」のＶＤＥＶは、３個のパス情報が対応付けられている。即ち、このＶＤＥＶ（＃３）にマッピングされる外部記憶デバイス(図１の４２)は、その内部に３つの経路を有する交代パス構造を備えているが、ＶＤＥＶには、この交代パス構造を認識してマッピングされている。これら３つの経路のいずれを通っても同一の記憶領域にアクセスできることが判明しているため、いずれか１つまたは２つの経路に障害等が発生した場合でも、残りの正常な経路を介して所望のデータにアクセスできる。図４に示すようなマッピングテーブルＴｍを採用することにより、第１の記憶制御装置２０内の１つ以上のＶＤＥＶに対し、１つまたは複数の外部の記憶デバイス４２をマッピングすることができる。

【００５０】

次に、図５を参照して、外部の記憶デバイス４２をＶＤＥＶ１０１にマッピングする方法の一例を説明する。図５は、マッピング時に、第１の記憶制御装置２０と第２の記憶制御装置４０との間で行われる処理の要部を示すタイムチャートである。まず、第１の記憶制御装置２０は、チャネルアダプタ２１のエクスターナルポート（２１Ａ）からスイッチ回路ＳＷを介して、第２の記憶制御装置４０にログインする（Ｓ１）。第２の記憶制御装置４０が、第１の記憶制御装置２０のログインに対して応答を返すことにより、ログインが完了する（Ｓ２）。次に、第１の記憶制御装置２０は、例えば、SCSI（Small Computer System Interface）規格で定められている照会コマンド（inquiryコマンド）を、第２の記憶制御装置４０に送信し、第２の記憶制御装置４０の有する記憶デバイス４２の詳細について応答を求める（Ｓ３）。これは既述のディスカバーの動作と同じである。

【００５１】

照会コマンドは、照会先の装置の種類及び構成を明らかにするために用いられるもので、照会先装置の有する階層を透過してその物理的構造を把握することができる。照会コマンドを使用することにより、第１の記憶制御装置２０は、例えば、装置名、デバイスタイプ、製造番号（プロダクトＩＤ）、ＬＤＥＶ番号、各種バージョン情報、ベンダＩＤ等の情報を第２の記憶制御装置４０から取得できる（Ｓ４）。第２の記憶制御装置４０は、問合わされた情報を第１の記憶制御装置２０に送信し、応答する（Ｓ５）。ディスカバリーで抽出された情報は共有メモリに保存されている。ＣＨＡはこの保存された情報を利用することによってＳ３～Ｓ５のステップを省略しても良い。

【００５２】

第１の記憶制御装置２０は、第２の記憶制御装置４０から取得した情報を、マッピングテーブルＴｍの所定箇所に登録する（Ｓ６）。次に、第１の記憶制御装置２０は、第２の

記憶制御装置４０から記憶デバイス４２の記憶容量を読み出す（Ｓ７）。第２の記憶制御装置４０は、第１の記憶制御装置２０からの問合せに対して、記憶デバイス４２の記憶容量を返信し（Ｓ８）、応答を返す（Ｓ９）。第１の記憶制御装置２０は、記憶デバイス４２の記憶容量をマッピングテーブルＴｍの所定箇所に登録する（Ｓ１０）。

【００５３】

以上の処理を行うことにより、マッピングテーブルＴｍを構築できる。第１の記憶制御装置２０のＶＤＥＶにマッピングされた外部の記憶デバイス４２（外部ＬＵＮ、即ち外部のＬＤＥＶ）との間でデータの入出力を行う場合は、後述する他のテーブルを参照してアドレス変換等を行う。

【００５４】

図６～図８を参照して、第１の記憶制御装置２０と第２の記憶制御装置４０との間のデータ入出力について説明する。まず最初に、データを書き込む場合について、図６及び図７に基づいて説明する。図６は、データ書込み時の処理を示す模式図である。図７は、図６中の処理の流れを各種テーブルとの関係で示す説明図である。

【００５５】

ホスト装置１０は、第１の記憶制御装置２０が提供する論理ボリューム（ＬＤＥＶ１０２）にデータを書き込むことができる。例えば、ＳＡＮの中に仮想的なＳＡＮサブネットを設定するゾーニングや、アクセス可能なＬＵＮのリストをホスト装置１０が保持するＬＵＮマスキングという手法により、ホスト装置１０を特定のＬＤＥＶ１０２に対してのみアクセスさせるように設定できる。

【００５６】

ホスト装置１０がデータを書き込もうとするＬＤＥＶ１０２が、ＶＤＥＶ１０１を介して内部の記憶デバイスである記憶デバイス３１に接続されている場合、通常の処理によってデータが書き込まれる。即ち、ホスト装置１０からのデータは、いったんキャッシュメモリ２４に格納され、キャッシュメモリ２４からディスクアダプタ２２を介して、所定の記憶デバイス３１の所定アドレスに格納される。この際、ディスクアダプタ２２は、論理的なアドレスを物理的なアドレスに変換する。また、RAID構成の場合、同一のデータが複数の記憶デバイス３１に記憶等される。

【００５７】

これに対し、ホスト装置１０が書き込もうとするＬＤＥＶ１０２が、ＶＤＥＶ１０１を介して外部の記憶デバイス４２に接続されている場合、図６に示すような流れでデータが書き込まれる。図６（ａ）は記憶階層を中心に示す流れ図であり、図６（ｂ）はキャッシュメモリ２４の使われ方を中心に示す流れ図である。

【００５８】

ホスト装置１０は、書込み先のＬＤＥＶ１０２を特定するＬＤＥＶ番号とこのＬＤＥＶ１０２にアクセスするための通信ポート２１Ａを特定するＷＷＮとを明示して、書込みコマンド（Write）を発行する（Ｓ２１）。第１の記憶制御装置２０は、ホスト装置１０からの書込みコマンドを受信すると、第２の記憶制御装置４０に送信するための書込みコマンドを生成し、第２の記憶制御装置４０に送信する（Ｓ２２）。第１の記憶制御装置２０は、ホスト装置１０から受信した書込みコマンド中の書込み先アドレス情報等を、外部ＬＤＥＶ４３に合わせて変更することにより、新たな書込みコマンドを生成する。

【００５９】

次に、ホスト装置１０は、書き込むべきデータを第１の記憶制御装置４０に送信する（Ｓ２３）。第１の記憶制御装置２０に受信されたデータは、ＬＤＥＶ１０２からＶＤＥＶ１０１を介して（Ｓ２４）、外部のＬＤＥＶ４３に転送される（Ｓ２６）。ここで、第１の記憶制御装置２０は、ホスト装置１０からのデータをキャッシュメモリ２４に格納した時点で、ホスト装置１０に対し書込み完了の応答（Good）を返す（Ｓ２５）。第２の記憶制御装置４０は、第１の記憶制御装置２０からデータを受信した時点で（あるいは記憶デバイス４２に書込みを終えた時点で）、書込み完了報告を第１の記憶制御装置２０に送信する（Ｓ２６）。即ち、第１の記憶制御装置２０がホスト装置１０に対して書込み完了を

報告する時期（Ｓ２５）と、実際にデータが記憶デバイス４２に記憶される時期とは相違する（非同期方式）。従って、ホスト装置１０は、実際にデータが記憶デバイス４２に格納される前にデータ書込み処理から解放され、別の処理を行うことができる。

【００６０】

図６（ｂ）のように、キャッシュメモリ２４には、多数のサブブロック２４Ａが設けられている。第１の記憶制御装置２０は、ホスト装置１０から指定された論理ブロックアドレスをサブブロックのアドレスに変換し、キャッシュメモリ２４の所定箇所にデータを格納する（Ｓ２４）。

【００６１】

図７を参照して、各種テーブルを利用してデータが変換される様子を説明する。図７の上部に示すように、ホスト装置１０は、所定の通信ポート２１Ａに対し、ＬＵＮ番号（ＬＵＮ＃）及び論理ブロックアドレス（ＬＢＡ）を指定してデータを送信する。第１の記憶制御装置２０は、ＬＤＥＶ１０２用に入力されたデータ（ＬＵＮ＃＋ＬＢＡ）を、図７（ａ）に示す第１の変換テーブルＴ１に基づいて、ＶＤＥＶ１０１用のデータに変換する。第１の変換テーブルＴ１は、内部のＬＵＮ１０３を指定するデータをＶＤＥＶ１０１用データに変換するための、ＬＵＮ－ＬＤＥＶ－ＶＤＥＶ変換テーブルである。このテーブルＴ１は、例えば、ＬＵＮ番号（ＬＵＮ＃）と、そのＬＵＮ１０３に対応するＬＤＥＶ１０２の番号（ＬＤＥＶ＃）及び最大スロット数と、ＬＤＥＶ１０２に対応するＶＤＥＶ１０１の番号（ＶＤＥＶ＃）及び最大スロット数等を対応付けることにより構成される。このテーブルＴ１を参照することにより、ホスト装置１０からのデータ（ＬＵＮ＃＋ＬＢＡ）は、ＶＤＥＶ１０１用のデータ（ＶＤＥＶ＃＋ＳＬＯＴ＃＋ＳＵＢＢＬＯＣＫ＃）に変換される。

【００６２】

次に、第１の記憶制御装置２０は、図７（ｂ）に示す第２の変換テーブルＴ２を参照して、ＶＤＥＶ１０１用のデータを、第２の記憶制御装置４０の外部ＬＵＮ（ＬＤＥＶ）用に送信して記憶させるためのデータに変換する。第２の変換テーブルＴ２には、例えば、ＶＤＥＶ１０１の番号（ＶＤＥＶ＃）と、そのＶＤＥＶ１０１からのデータを第２の記憶制御装置４０に送信するためのイニシエータポートの番号と、データ転送先の通信ポート４１を特定するためのＷＷＮと、その通信ポートを介してアクセス可能なＬＵＮ番号とが対応付けられている。この第２の変換テーブルＴ２に基づいて、第１の記憶制御装置２０は、記憶させるべきデータの宛先情報を、イニシエータポート（ターゲットポート）番号＃＋ＷＷＮ＋ＬＵＮ＃＋ＬＢＡの形式に変換する。このように宛先情報が変更されたデータは、指定されたイニシエータポートから通信ネットワークＣＮ１を介して、指定された通信ポート４１に到達する。そして、データは、指定されたＬＵＮ４３でアクセス可能なＬＤＥＶの所定の場所に格納される。ＬＤＥＶは、複数の記憶デバイス４２上に仮想的に構築されているので、データのアドレスは物理アドレスに変換されて、所定のディスクの所定アドレスに格納される。

【００６３】

図７（ｃ）は、別の第２の変換テーブルＴ２ａを示す。この変換テーブルＴ２ａは、外部記憶デバイス４２に由来するＶＤＥＶ１０１に、ストライプやRAIDを適用する場合に使用される。変換テーブルＴ２ａは、ＶＤＥＶ番号（ＶＤＥＶ＃）と、ストライプサイズと、RAIDレベルと、第２の記憶制御装置４０を識別するための番号（ＳＳ＃（ストレージシステム番号））と、イニシエータポート番号と、通信ポート４１のＷＷＮ及びＬＵＮ４３の番号とを対応付けることにより構成されている。図７（ｃ）に示す例では、１つのＶＤＥＶ１０１は、ＳＳ＃（１，４，６，７）で特定される合計４つの外部記憶制御装置を利用してRAID１を構成する。また、ＳＳ＃１に割り当てられている３個のＬＵＮ（＃０，＃０，＃４）は、同一デバイス（ＬＤＥＶ＃）に設定されている。なお、ＬＵＮ＃０のボリュームは、２個のアクセスデータパスを有する交代パス構造を備えている。このように、本実施例では、外部に存在する複数の論理ボリューム（ＬＤＥＶ）からＶＤＥＶ１０１を構成することにより、ストライピングやRAID等の機能を追加した上でホスト装置１０に提

供することができる。

【００６４】

図８を参照して、第２の記憶制御装置４０のＬＤＥＶからデータを読み出す場合の流れを説明する。まず、ホスト装置１０は、通信ポート２１Ａを指定して第１の記憶制御装置２０にデータの読み出しコマンドを送信する（Ｓ３１）。第１の記憶制御装置２０は、読み出しコマンドを受信すると、要求されたデータを第２の記憶制御装置４０から読み出すべく、読み出しコマンドを生成する。第１の記憶制御装置２０は、生成した読み出しコマンドを第２の記憶制御装置４０に送信する（Ｓ３２）。第２の記憶制御装置４０は、第１の記憶制御装置２０から受信した読み出しコマンドに応じて、要求されたデータを記憶デバイス４２から読み出して、第１の記憶制御装置２０に送信し（Ｓ３３）、正常に読み出しが完了した旨を報告する（Ｓ３５）。第１の記憶制御装置２０は、図８（ｂ）に示すように、第２の記憶制御装置４０から受信したデータを、キャッシュメモリ２４の所定の場所に格納させる（Ｓ３４）。

【００６５】

第１の記憶制御装置２０は、キャッシュメモリ２４に格納されたデータを読み出し、アドレス変換を行った後、ＬＵＮ１０３等を介してホスト装置１０にデータを送信し（Ｓ３６）、読み出し完了報告を行う（Ｓ３７）。これらデータ読み出し時の一連の処理では、図７と共に述べた変換操作が逆向きで行われる。

【００６６】

図８では、ホスト装置１０からの要求に応じて、第１の制御装置のチャネルアダプタ２１が第２の記憶制御装置４０からデータを読み出し、仮想ボリュームを構成するキャッシュメモリ２４に保存することを示している。符号１０３は、仮想ボリュームを構成するＬＵＮを示し、キャッシュメモリ２４は仮想ボリュームの実記憶領域に相当する。外部ＬＵＮは後述のリストアデータが記憶された論理ボリュームである。

【００６７】

次に、本願発明者が今回新たに認知した、本願発明を成立させる契機となった課題について説明する。図９は、第２の記憶制御装置４０が第１の記憶制御装置２０に接続した記憶制御システムのブロックである。図９（１）はリストア前のデータの記憶状況を示したブロック図であり、（２）はリストア後のデータの記憶状況を示すブロック図である。第１の記憶制御システム２０内には、実論理ボリュームＬＵ＃ａと仮想ボリュームＬＵ＃ｂとが、特開２００１－２１６１８５号公報に記載のコピーペアの関係を形成可能に構成されている。

【００６８】

ＬＵ＃ａが正ボリューム番号によって識別される正ボリュームであり、ＬＵ＃ｂが副ボリューム番号で識別される副ボリュームである。第２の記憶制御装置４０も同様に、互いにコピーペアの関係になる実論理ボリュームＬＵ＃ｃ(正ボリューム)と実論理ボリューム＃ｄ(副ボリューム)とを備えている。これら正ボリューム及び副ボリュームとがコピーペア形成中であれば、正ボリュームのデータと副ボリュームの一方がコピー元、他方がコピー先にとなり、両者間で全コピーが行われる。

【００６９】

第１の記憶制御装置の仮想ボリュームＬＵ＃ｂは図２で説明したように、論理ボリュームＬＵ＃ｃにマッピングされている。論理ボリュームＬＵはＬＤＥＶであっても良い。実論理ボリュームＬＵ＃ａ，ＬＵ＃ｃ，ＬＵ＃ｄのそれぞれは、ディスクドライブの物理的な記憶領域とキャッシュメモリに対応しており、データのＲＥＡＤ命令及びＷＲＩＴＥ命令の実行に伴うデータの交換はキャッシュメモリを介してＨＤＤなどの記憶デバイスとの間で行われる。

【００７０】

仮想ボリュームＬＵ＃ｂは対応する記憶デバイスを持たず、キャッシュメモリのみから構成されている。論理ボリュームＬＵ＃ｃのデータに仮想ボリュームＬＵ＃ｂを介してホスト＃ｂがアクセスができる。各論理ボリュームの上段がキャッシュメモリに記憶された

データに対応し、下段がディスクドライブ内のデータに対応している。

【００７１】

図９に示すシステムによれば、第１の記憶制御装置に第２の記憶制御装置が接続している関係において、正ボリュームから副ボリュームにデータをコピーするコピーペアの機能を利用して、データの世代管理が可能となる。すなわち、ホスト＃ａがアクセス可能な実論理ボリュームＬＵ＃ａには最新世代のデータを記憶し、ホスト＃ｂがアクセス可能な仮想ボリューム＃ｂ及び実論理ボリューム＃ｃには前の世代のデータを記憶し、ホスト＃ｂがアクセス可能な実論理ボリュームＬＵ＃ｄにはさらに前の世代のデータを記憶することができる。例えば、最新世代のデータをオンラインデータとすると、一つ前の世代のデータを過去データ１（例えば一日前のデータ）とし、さらに前の世代のデータを過去データ２（例えば二日前のデータ）として、図９の記憶制御システムでデータの世代管理をすることができる。

【００７２】

本発明に於けるリストアは、第２の記憶制御装置の論理ボリュームＬＵ＃ｄのデータを論理ボリュームＬＵ＃ｃに記憶させることをいう。ホスト＃ｂは仮想ボリュームＬＵ＃ｂにアクセスすることによって、論理ボリュームＬＵ＃ｃに記憶されたリストアデータにアクセスすることができる。リストによって、記憶制御装置２０のデータを記憶制御装置４０の記憶資源に記憶された過去データに基づいて復旧することができる。

【００７３】

リストアの処理について説明する。既述のとおり、図９の（１）はリストア前の記憶制御システムのデータの記憶状態を示し、（２）はリストア実行後の状態を示している。第２の記憶制御装置４０のＤＫＡが、ホスト＃ｂからリストア命令を受けると、論理ボリュームＬＵ＃ｄの記憶デバイス９００からリストアの対象となるデータ(リストアデータ)「ｄｅｆ」を読み込み、これを論理ボリュームＬＵ＃ｄのキャッシュメモリ９０２の記憶領域に保存する（Ｓ９０）。ＣＨＡ９０Ａはこのキャッシュメモリ９０２の記憶領域のデータを論理ボリュームＬＵ＃ｃのキャッシュメモリ９０４の記憶領域に記憶する（Ｓ９１）。そして、この記憶領域のデータ「ｄｅｆ」をＤＫＡが論理ボリュームＬＵ＃ｃの記憶デバイス９０６の記憶領域に記憶する（Ｓ９２）。

【００７４】

一方、第１の記憶制御装置２０のＣＨＡ９０Ｂは、ホスト＃ｂからのリストア命令を実行し、リストアデータを保存するキャッシュメモリ９０８のブロック領域を設定し、この設定情報を第２の記憶制御装置４０のＣＨＡ９０Ａに転送する(Ｓ９４)。ＣＨＡ９０Ａはリストアデータが記録されているキャッシュメモリ９０４の記憶領域のデータ「ｄｅｆ」を読み込み（Ｓ９５）、第１の記憶制御装置４０のＣＨＡ９０Ｂに渡す（Ｓ９６）。ＣＨＡ９０Ｂは予め決められたキャッシュメモリ９０８のブロック記憶領域にこのデータ「ｄｅｆ」を保存する（Ｓ９３）。

【００７５】

なお、図９（２）には示されていないが、論理ボリュームＬＵ＃ａと仮想ボリューム＃ｂとがコピーペアを形成すれば、仮想ボリュームＬＵ＃ｂに対応するキャッシュメモリ９０８の記憶領域のデータが論理ボリュームＬＵ＃ａの記憶領域(キャッシュメモリ及びＨＤＤ)にコピーされる。

【００７６】

いま、第２の記憶制御装置４０がリストア命令を実行すると、図９（２）に示すように、論理ボリュームＬＵ＃ｃのデータは「ａｂｃ」から「ｄｅｆ」になる。この論理ボリュームＬＵ＃ｃにマッピングされている仮想ボリュームＬＵ＃ｂには「ｄｅｆ」が転送される。しかしながら、仮想ボリュームに対応するキャッシュメモリの記憶領域にはリストア前の「ａｂｃ」のデータが残っているので、ホスト＃ｂにはリストアデータである「ｄｅｆ」ではなく、リストア前のデータである「ａｂｃ」が見てしまう。

【００７７】

さらに、仮想ボリュームＬＵ＃ｂとコピーペアの関係にある正論理ボリュームＬＵ＃ａ

にも「ｄｅｆ」ではなく「ａｂｃ」をコピーしてしまうことになる。これでは、第２の記憶制御装置４０のリストアデータに、ホスト＃ｂが正しくアクセスできない問題がある。

【００７８】

そこで、第１の記憶制御装置１０は、ホスト＃ｂのパージコマンドを受けて、リストア命令の実行の際に仮想ボリュームＬＵ＃ｂに対応するキャッシュメモリのデータ「ａｂｃ」をパージ（「無効化」或いは「開放」）することとした（Ｓ９７）。ここで、パージとは、仮想ボリュームとキャッシュメモリの記憶領域との対応関係をクリアするための処理に相当する。この処理が終了した後、或いはこの処理に同期して、第１の記憶制御装置１０は、キャッシュメモリの記憶領域にリストアデータ「ｄｅｆ」を記憶することとした。

【００７９】

図１０は、仮想ボリュームのディレクトリ構造とキャッシュメモリとの対応関係を示すブロック図である。仮想ボリュームＬＵ＃ｂは複数の管理領域１００から構成されていて、１００個のＬＢＡ毎に１個の管理領域１００が割り当っている。一つの管理領域はキャッシュメモリの記憶領域を指す番号(キャッシュ領域＃)１００Ａと、キャッシュ領域＃の中のデータが有効かどうかを示す有効ビットマップ１００Ｂとから構成される。

【００８０】

図１０は、ＬＢＡ＃１００～ＬＢＡ＃１９９にキャッシュメモリの記憶領域（キャッシュ領域）＃１の各ブロックが対応していることを示している。キャッシュ領域＃は、各ＬＢＡに対応したブロックに分割されており、各ブロックに各ＬＢＡに対応するデータが置かれる。ビットマップの各ビットは、一つのＬＢＡ及びこれに対応するキャッシュ領域のブロック単位に対応しており、有効ビットが「０」の場合には、キャッシュ領域の該当するブロック単位のデータが無効であることを示し、有効ビットが「１」の場合には、対応ブロック単位にあるデータが有効であることを示している。有効ビットが「０」の場合には、ホストから対応するＬＢＡにリード要求が来た場合には、第２の記憶制御装置の論理ボリュームのデータをリードする必要がある。

【００８１】

キャッシュ領域＃１においては、ＬＢＡ＃１００とＬＢＡ＃１０１に対応するブロック単位のデータが有効であり、ＬＢＡ＃１０２，１０３に対応するブロック単位のデータが無効である事を示している。管理領域１００の［ＮＵＬＬ］は、キャッシュ領域に対応する部分がないことを示している。仮想ボリュームの管理領域に割り当っていないキャッシュ領域は、フリーキューとして共有メモリ(図１の２５)の特定記憶領域に記憶されている。

【００８２】

ホストからのリード命令が、キャッシュ領域の割り当てがない、仮想ボリュームの管理領域に来た場合には、第１の制御装置のＣＨＡは共有メモリのフリーキューの中からキャッシュ領域の情報を取得し、この情報を仮想ボリュームの管理領域に割り当てる。図１０に示す仮想ボリュームのディレクトリ構造を規定する制御情報は共有メモリに記憶される。なお、図１０は仮想ボリュームについて説明したが、実論理ボリュームでも同様である。

【００８３】

図１１は、この実施形態におけるパージ処理を説明するフローチャートである。図９（２）のホスト＃ｂがパージ処理の対象となる仮想ボリュームを指定し、この仮想ボリュームの第１の管理領域に対して(Ｓ１１００)、パージ処理を実行する（Ｓ１１０２－Ｓ１１０６）。

【００８４】

ホストが第１の記憶制御装置４０にパージコマンドを発行すると、ＣＨＡ９０Ｂは、共有メモリの仮想ボリュームのディレクトリ構造(図１０)をリードし、指定された管理領域にキャッシュメモリの記憶領域の情報が設定されているかを判定する(Ｓ１１０２)。ＣＨＡ９０Ｂは、これが肯定された場合には、割り当てられたキャッシュ領域をフリーキューに戻して、このキャッシュ領域に「ＮＵＬＬ」を設定する（Ｓ１１０４）。Ｓ１１０２の

判定が否定された場合及びＳ１１０４の処理に継いで、ＣＨＡは全管理領域についてＳ１１０２及びＳ１１０４の処理が完了したか否かを判定する。それが肯定された場合は、ルーチンを終了する。否定判定の場合には、次の管理領域について、Ｓ１１０２、Ｓ１１０４の処理を継続する（Ｓ１１０８）。これにより、仮想ボリュームの全管理領域に対してキャッシュメモリの記憶領域のデータがパージされたことになり、リストアされる前の、キャッシュメモリのデータが仮想ボリュームのディレクトリ構造を介してホストにアクセスされることが無いようにしている。

【００８５】

このパージが完了されると、第１の記憶制御装置のＣＨＡ９０Ｂ（図９）はパージ完了をホスト＃ｂに通知する。ホストは第２の記憶制御装置のＣＨＡ９０Ａにリストア要求を発行する。リストアの命令は、コマンドデバイス（図２の１６６）を経由して第１の制御装置のＣＨＡ９０Ｂが第２の制御装置のＣＨＡ９０Ａにホスト＃ｂを介することなく送ることもできる。リストア命令を受けた第２の記憶装置は、図９に示すように論理ボリュームＬＵ＃ｄとＬＵ＃ｃとのコピーペアを形成して、論理ボリュームＬＵ＃ｄのデータを論理ボリュームＬＵ＃ｃにコピーする（リストア）。ホストからアクセスのあった仮想ボリュームの管理領域には、フリーキューの中から適宜キャッシュメモリの領域が選択されて割り当てられ、このキャッシュメモリの領域にリストアデータが記憶される。論理ボリュームＬＵ＃ｃは仮想ボリュームＬＵ＃ｂにマッピングされているので、論理ボリュームＬＵ＃ｃのキャッシュメモリのリストアデータが、割り当てられたキャッシュ領域に転送される。したがって、ホスト＃ｂがＣＨＡ９０Ｂを介して、仮想ボリュームの管理領域にアクセスすることにより、ホスト＃ｂは、この管理領域に対応するキャッシュメモリの記憶領域にあるリストアデータにアクセスすることができる。

【００８６】

仮想ボリュームの管理領域に［ＮＵＬＬ］を設定することにより、キャッシュメモリのリストア前のデータをパージすることができる。なお、パージ処理の他の例として、有効ビット（図１０の１００Ｂ）に「０」を設定するようにしてもよい。さらに、キャッシュのパージが終了する前に、第２の記憶制御装置はリストアを開始してもよい。ただし、仮想ボリュームの全ての管理領域についてキャッシュメモリのパージが終了しないと、ホスト＃ｂは仮想ボリュームＬＵ＃ｂのリストアデータにアクセスすることができない。今後論理ボリュームの容量、キャッシュ容量がますます大きくなっていくため、パージに要する時間も長くなり、その結果リストアしたデータをホストが利用できるようになるまでに時間がかかってしまう。

【００８７】

そこで、第２の実施形態では、ホストがアクセスした、仮想ボリュームに対応するキャッシュメモリの記憶領域のデータがリストアデータを反映したものであるか否かを第１の記憶制御装置が判断し、リストアデータを反映していない場合には、このキャッシュ領域のデータをパージし、ついでこのキャッシュ領域にリストアデータを記憶するようにして、全ての管理領域についてパージを待つことなく、ホスト装置がリストアデータにアクセスできるようにした。

【００８８】

このことを可能にするための、仮想ボリュームのディレクトリ構造が図１２に示されている。仮想ボリュームに相当するＬＵ毎に全体世代番号１２０が設定され、各管理領域には図１０に示す、キャッシュ領域１００Ａ、有効ビット１００Ｂの他、個別の世代番号１２０Ａも設定される。図１２では、仮想ボリュームのキャッシュ領域＃１がキャッシュメモリ２４のキャッシュ領域＃１に、仮想ボリュームのキャッシュ領域＃５がキャッシュメモリ２４のキャッシュ領域＃５に対応している。

【００８９】

図１３はこの実施形態のパージ処理のためのフローチャートを示すものであり、図９（２）に示す、ホスト＃ｂから記憶制御装置２０にパージ要求が出力されると、ＣＨＡ９０Ｂは共有メモリの制御情報にアクセスして仮想ボリュームを指定し、仮想ボリュームのデ

ィレクトリ構造の全体世代番号１２０の値を１インクリメントする（Ｓ１３００）。ついでＣＨＡ９０ＢはホストＨｂに正常終了を通知する。ホストＨｂの命令を受けて、記憶制御装置４０は二つの内部論理ボリューム間でシャドウイメージを利用したリストア処理を開始する。

【００９０】

ホストＨｂがこの正常終了を受けると、図１４に示す、リストアデータへのアクセス処理（Ｉ／Ｏ処理）がスタートする。ホストＨｂが仮想ボリュームＬＵＨｂ(図９（２）参照)へＩ／Ｏを発行すると、第１の記憶制御装置のＣＨＡ９０ＢはホストＨｂからＩ／Ｏの要求を受けた仮想ボリュームのディレクトリ構造を共有メモリから参照し、Ｉ／Ｏ要求を受けたＬＢＡのキャッシュ領域（図１２）を取得する（Ｓ１４００）。ＣＨＡはキャッシュ領域の制御データがＮＵＬＬであるか否かをチェックする（Ｓ１４０１）。これが肯定されると、ＣＨＡはＬＢＡが属する管理領域にキャッシュメモリのフリーキューのキャッシュ領域を割り当てる。さらに有効ビット（図１２の１００Ｂ）をオフし、世代番号に代表世代番号を設定する（Ｓ１４０２）。

【００９１】

ステップＳ１４０１において、キャッシュ領域を示す情報がＮＵＬＬでない場合には、ＣＨＡは、Ｉ／Ｏ要求を受けたＬＢＡが属する管理領域の世代番号と全体世代番号とを比較する（Ｓ１４０４）。この比較結果が否の場合には管理領域の有効ビットがオフされる（Ｓ１４０６）。これが肯定されている場合には、Ｓ１４０６をスキップしてＳ１４０８にジャンプする。

【００９２】

Ｓ１４０８では、ＣＨＡが、ホストからの命令がWRITE命令又はREAD命令かを判定し、後者の場合はＳ１４１０に移行する。ＣＨＡはＬＵＨｃからリストアデータを読み出して、ＬＢＡで指定される、仮想ボリュームＬＵＨｂにおける管理領域上のキャッシュ領域Ｈに該当するキャッシュメモリの記憶領域にリストアデータを書き込む。次いで，ＣＨＡは管理領域の有効ビットをオンし（Ｓ１４１２）、ホストにリストアデータを転送する（Ｓ１４１４）。

【００９３】

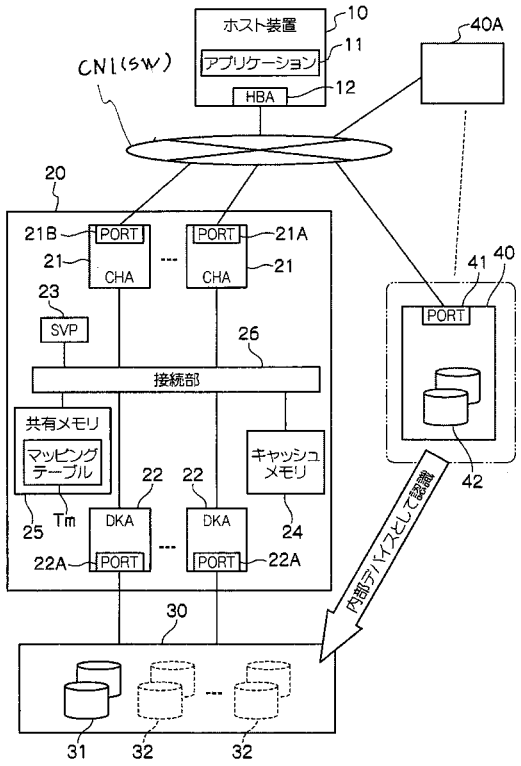
Ｓ１４０８で、ホストからの命令がWRITE命令の場合には、ＣＨＡはホストからライトデータを受領して、Ｉ／Ｏ要求を受けたＬＢＡのキャッシュ領域に該当するキャッシュメモリ上の記憶領域にライトデータを書き込んで(Ｓ１４１６)、有効ビットをオンする(Ｓ１４１８)。

【００９４】

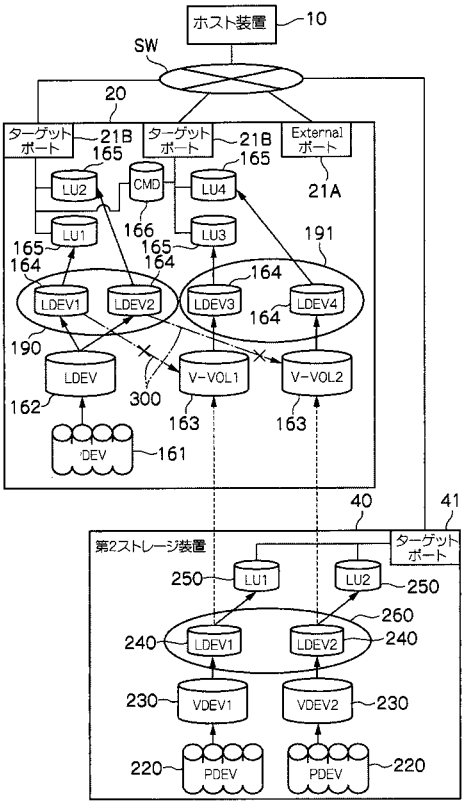
この一連の処理において、第１の記憶制御装置のＣＨＡは、ホストからのＩ／Ｏ要求を受けたＬＢＡに対応する、仮想ボリュームの管理領域の世代番号と全体世代番号とを比較し（Ｓ１４０４）、これが一致する場合には、キャッシュメモリの記憶領域に正しいリストアデータをアップロードできる(Ｓ１４１０)。一方、この世代番号が一致していない場合には、キャッシュメモリの記憶領域のデータはリストアデータに対応しない不適切なデータであるので、第１の記憶制御装置のＣＨＡは有効ビットをオフしてなる、キャッシュメモリのデータのパージを行い（Ｓ１４０６）、次いで、仮想ボリュームの管理領域にリストアデータを設定する。この実施形態では、第１の制御装置の制御部がキャッシュメモリ内のデータがパージを必要とするリストア前のデータか、リストアデータであってホストをアクセスさせても良いデータかを識別し、パージが必要なデータである場合には、外部論理ボリュームからリストアデータをリードして、キャッシュメモリの記憶領域のリストア前のデータをリストアデータでリフレッシュする。この実施形態では、仮想ボリュームの全管理領域に対してキャッシュメモリのパージを行わず、ホストがアクセスした管理領域について、キャッシュメモリのデータのパージを行えばすむ為、ホストがリストアデータにアクセスするために要する時間を短縮化できるという効果がある。また、この実施形態はパージ処理とリストアデータの転送とを並列処理するために、ホストがリストアデータを利用できるまでに要する時間を短縮化できる。

【００９５】
なお、既述の実施形態では、第２の記憶制御装置内の副ボリュームから正ボリュームに全コピーしてデータのリストアを行っていたが、論理ボリュームＬＵ＃ｃ(図９)にホスト＃ｃが接続され、このホストに接続したテープデバイスなどの記憶資源からデータのリストアを行う場合にも本発明が適用される。また、仮想ボリュームに対するパージ処理は仮想ボリューム全体、管理領域単位、或いはＬＢＡ（ロジカルブロックアドレス）の単位で行うことができる。
【図面の簡単な説明】
【００９６】
【図１】本発明の実施例に係わる記憶システムの全体構成を示すブロック図である。
【図２】記憶システムの論理的構成の概要を示す模式図である。
【図３】仮想論理ボリュームの設定動作の流れを示すフローである。
【図４】マッピングテーブルの概要を示す説明図である。
【図５】マッピングテーブルを構築するための処理の流れを示すフローである。
【図６】内部ボリュームとして仮想化される外部の記憶デバイスにデータを書き込む場合の概念図である。
【図７】書込みデータのアドレス変換の様子を模式的に示す説明図である。
【図８】内部ボリュームとして仮想化される外部の記憶デバイスからデータを読み出す場合の概念図である。
【図９】キャッシュメモリのパージ処理を説明するための、記憶制御システムのブロック図であり、（１）はリストア前のデータの記憶状態を、（２）はリストア後のデータの記憶状態を説明するブロック図である。
【図１０】第１の実施形態における、仮想ボリュームのディレクトリ構造と、このディレクトリ構造に対応するキャッシュメモリの記憶構造とのブロック図である。
【図１１】第１の実施形態のパージ処理を説明するフローチャートである。
【図１２】第２の実施形態における、仮想ボリュームのディレクトリ構造と、このディレクトリ構造に対応するキャッシュメモリの記憶構造とのブロック図である。
【図１３】第２の実施形態における、パージ処理のフローチャートである。
【図１４】ホストがリストアデータにアクセスする事を示すフローチャートである。。
【符号の説明】
【００９７】
１０…ホスト装置、１１…アプリケーションプログラム、１２…アダプタ、２０…第１の記憶制御装置、２１，９０Ａ，９０Ｂ…チャネルアダプタ、２１Ａ…通信ポート、２２…ディスクアダプタ、２２Ａ…通信ポート、２３…コントロールユニット、２４…キャッシュメモリ、２４Ａ…サブブロック、２５…共有メモリ、２６…接続部、３０…記憶装置、３１…記憶デバイス、３２，ＬＵ＃ｂ…記憶デバイス（仮想化された内部記憶デバイス/仮想ボリューム）、４０…第２の記憶制御装置、４１…各通信ポート、４２…各記憶デバイス、ＬＵ＃ａ，ＬＵ＃ｃ，ＬＵ＃ｄ…実論理ボリューム、＃ａ，＃ｂ…ホスト

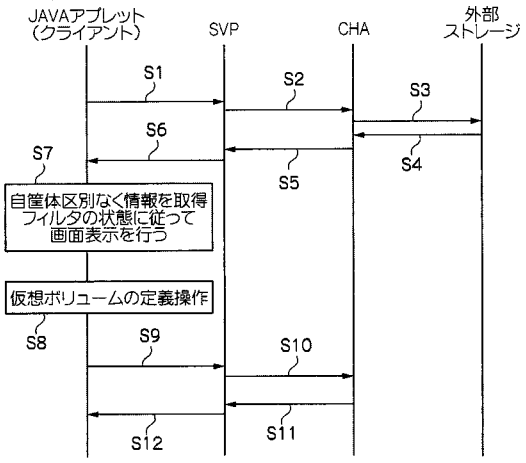
【図１】



【図２】



【図３】

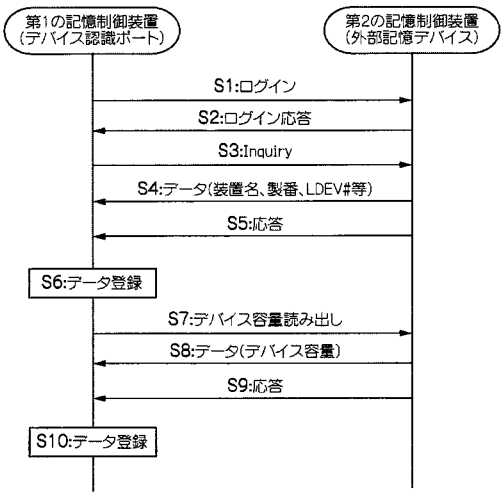


【図４】

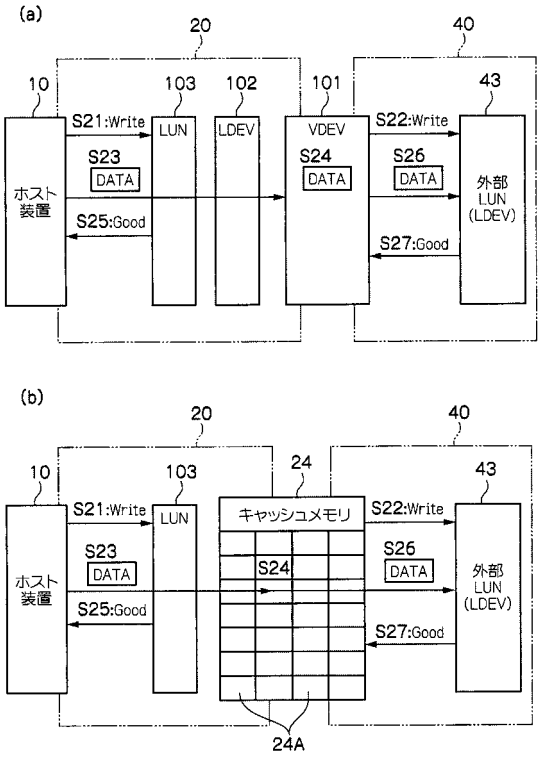
Tm

外部デバイス情報							交代パス
VDEV	デバイス識別情報	容量(KB)	デバイス種別	パス情報 WWN	パス情報 LUN		
0	DRFGTFNEIEK	657,456	DISK	0xAABBCCDD	0		
1	ADRFGTFNEIE	89,854	DISK	0xAABBEEFF	3		
2	GGRRFFDDERT	-	TAPE	0x445566AAB	5	6	0x77DE12345
3	AABBCCDDEE	5,544223	DISK	0x77DE12345	3	5	0x377DE7890

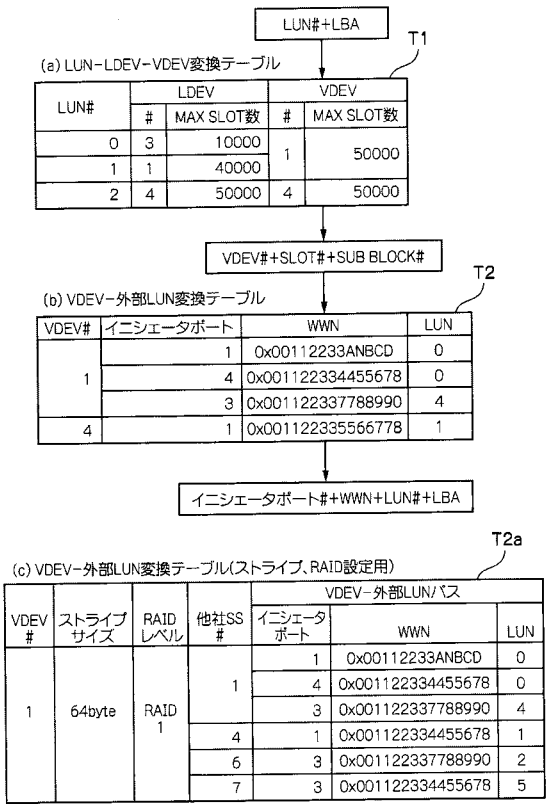
【図５】


第1の記憶制御装置
(デバイス認識ポート)
第2の記憶制御装置
(外部記憶デバイス)
S1:ログイン
S2:ログイン応答
S3:Inquiry
S4:データ(装置名、製番、LDEV#等)
S5:応答
S6:データ登録
S7:デバイス容量読み出し
S8:データ(デバイス容量)
S9:応答
S10:データ登録

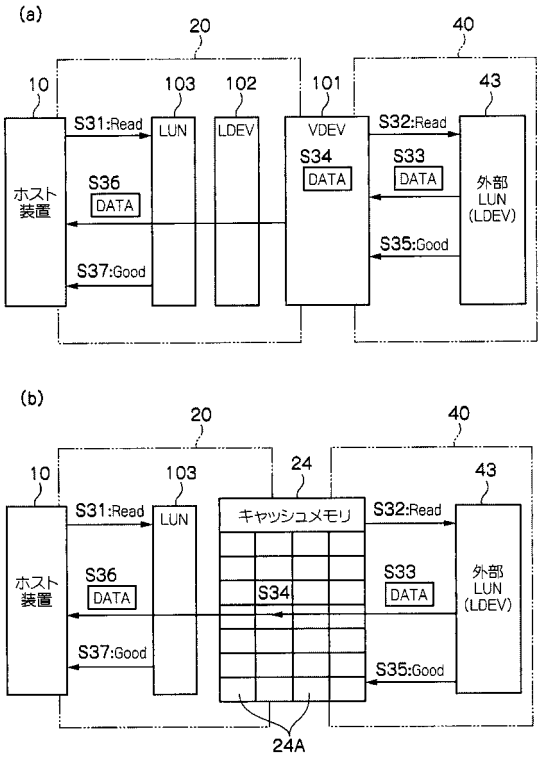
【図６】


(a)
20
40
10
103
102
101
43
S21:Write
LUN
LDEV
VDEV
S22:Write
S23
DATA
S24
DATA
S26
DATA
ホスト装置
外部LUN(LDEV)
S25:Good
S27:Good
(b)
20
40
10
103
24
43
S21:Write
LUN
キャッシュメモリ
S22:Write
S23
DATA
S24
S26
DATA
ホスト装置
外部LUN(LDEV)
S25:Good
S27:Good
24A

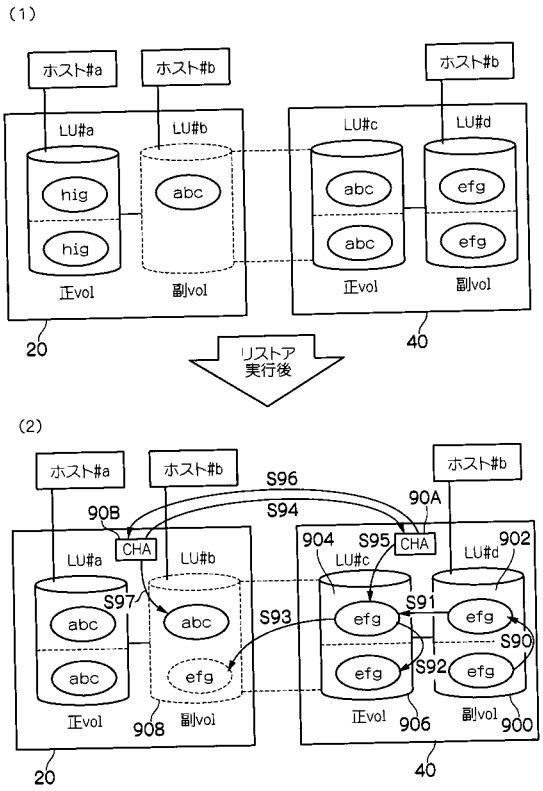
【図７】


LUN#+LBA
T1
(a) LUN-LDEV-VDEV変換テーブル
LUN#
LDEV
VDEV
#
MAX SLOT数
#
MAX SLOT数
0 3 10000 1 50000
1 1 40000
2 4 50000 4 50000
VDEV#+SLOT#+SUB BLOCK#
T2
(b) VDEV-外部LUN変換テーブル
VDEV#
イニシエータポート
WWN
LUN
1 1 0x00112233ANBCD 0
4 0x001122334455678 0
3 0x001122337788990 4
4 1 0x001122335566778 1
イニシエータポート#+WWN+LUN#+LBA
T2a
(c) VDEV-外部LUN変換テーブル(ストライプ、RAID設定用)
VDEV#
ストライプサイズ
RAIDレベル
他社SS#
VDEV-外部LUNパス
イニシエータポート
WWN
LUN
1 64byte RAID 1 1 1 0x00112233ANBCD 0
4 0x001122334455678 0
3 0x001122337788990 4
4 1 0x001122334455678 1
6 3 0x001122337788990 2
7 3 0x001122334455678 5

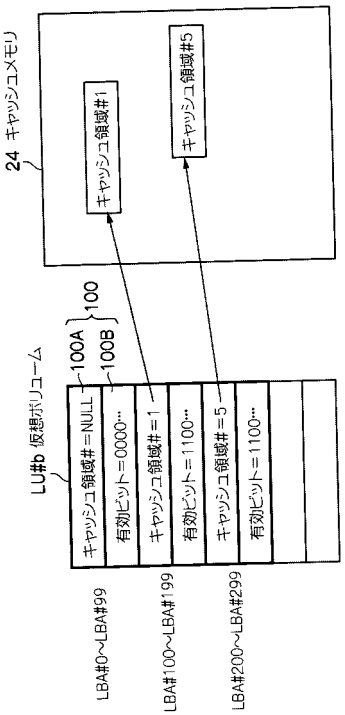
【図８】


(a)
20
40
10
103
102
101
43
S31:Read
LUN
LDEV
VDEV
S32:Read
S34
DATA
S33
DATA
S36
DATA
ホスト装置
外部LUN(LDEV)
S35:Good
S37:Good
(b)
20
40
10
103
24
43
S31:Read
LUN
キャッシュメモリ
S32:Read
S36
DATA
S34
S33
DATA
ホスト装置
外部LUN(LDEV)
S37:Good
S35:Good
24A

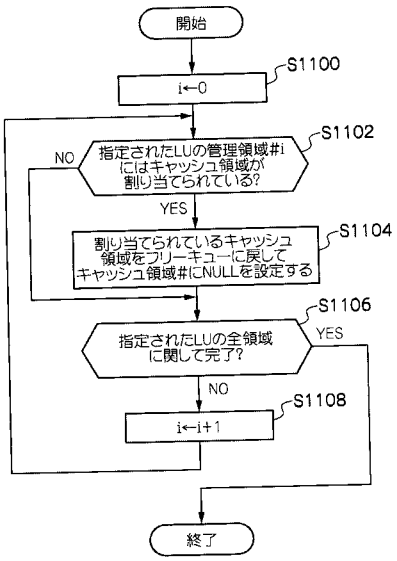
【図９】



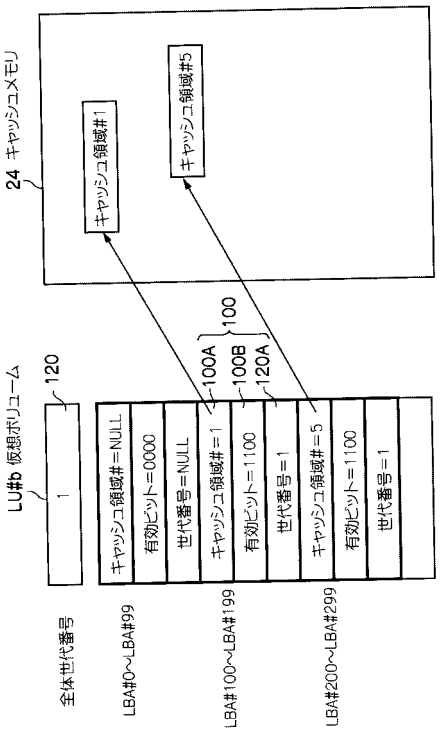
【図１０】



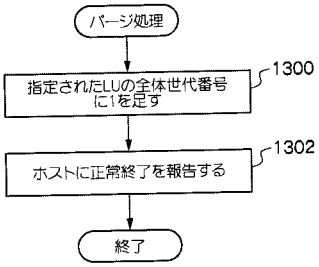
【図１１】



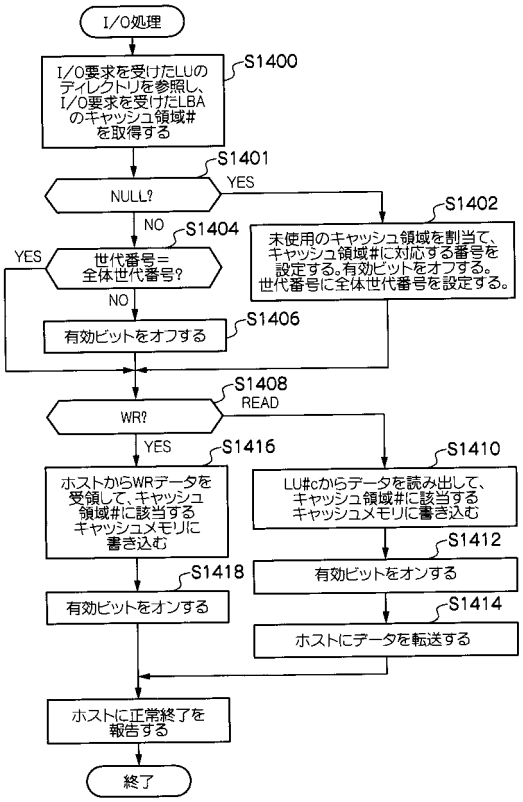
【図１２】



【図１３】


パージ処理
指定されたLUの全体世代番号に1を足す
1300
ホストに正常終了を報告する
1302
終了

【図１４】


I/O処理
I/O要求を受けたLUのディレクトリを参照し、I/O要求を受けたLBAのキャッシュ領域#を取得する
S1400
NULL?
S1401
YES
NO
S1402
未使用のキャッシュ領域を割当て、キャッシュ領域#に対応する番号を設定する。有効ビットをオフする。世代番号に全体世代番号を設定する。
世代番号＝全体世代番号?
S1404
YES
NO
有効ビットをオフする
S1406
WR?
S1408
READ
YES
ホストからWRデータを受領して、キャッシュ領域#に該当するキャッシュメモリに書き込む
S1416
LU#cからデータを読み出して、キャッシュ領域#に該当するキャッシュメモリに書き込む
S1410
有効ビットをオンする
S1412
有効ビットをオンする
S1418
ホストにデータを転送する
S1414
ホストに正常終了を報告する
終了

フロントページの続き

(51)Int.Cl.	FI		テーマコード（参考）
	Ｇ０６Ｆ 12/08	５３１Ｚ	
	Ｇ０６Ｆ 12/08	５５７	
	Ｇ０６Ｆ 12/08	５５９Ｚ	
	Ｇ０６Ｆ 13/12	３４０Ｂ	

Ｆターム(参考) 5B065 BA01 CA12 CC03 CC08 CE12 CH01 EA02 EA25