

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号  
特許第4887893号  
(P4887893)

(45) 発行日 平成24年2月29日 (2012. 2. 29)

(24) 登録日 平成23年12月22日 (2011. 12. 22)

(51) Int. Cl.

F I

G O 6 F 12/00 (2006. 01)

G O 6 F 3/06 (2006. 01)

G O 6 F 12/00 5 3 1 D

G O 6 F 12/00 5 3 3 J

G O 6 F 3/06 3 0 4 F

請求項の数 8 (全 42 頁)

(21) 出願番号	特願2006-121541 (P2006-121541)	(73) 特許権者	000005108
(22) 出願日	平成18年4月26日 (2006. 4. 26)		株式会社日立製作所
(65) 公開番号	特開2007-293652 (P2007-293652A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成19年11月8日 (2007. 11. 8)	(74) 代理人	100100310
審査請求日	平成20年12月25日 (2008. 12. 25)		弁理士 井上 学
		(72) 発明者	平岩 友理
			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	牧 晋広
			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

最終頁に続く

(54) 【発明の名称】 計算機システム及び計算機システムの制御方法

(57) 【特許請求の範囲】

【請求項 1】

ホストコンピュータと、  
第一のグループを構成する複数の第一論理ボリュームを有し、前記ホストコンピュータに接続する第一の記憶システムと、  
前記複数の第一論理ボリュームに対応した複数の第二論理ボリュームを有し、同期リモートコピー方式を用いて、前記複数の第一論理ボリュームのデータの複製を前記複数の第二論理ボリュームに格納する、第二の記憶システムと、  
前記第一のグループに対応する第三のグループを構成する複数の第三論理ボリュームを有し、非同期リモートコピー方式を用いて、前記第一のグループのデータの複製を前記第三のグループを構成する前記複数の第三論理ボリュームに格納する、第三の記憶システムと、  
を有し、  
前記第二の記憶システムは、同期リモートコピーを行う前記複数の第一論理ボリュームと前記複数の第二論理ボリュームとの第一のペアを管理し、  
前記第三の記憶システムは、非同期リモートコピーを行う前記第一のグループと前記第三のグループとの第二のペアを管理し、  
前記第二の記憶システムは、  
さらに、前記複数の第二論理ボリュームを含む第二のグループと、前記第三のグループとの第三のペアを管理し、

前記第一及び第二のペアにおけるリモートコピーが行われている間に、前記ホストコンピュータより、前記第二のグループのデータのうち前記第三のグループが格納していないデータである差分データを前記第三のグループに転送する、非同期リモートコピーの予約指示を受信すると、前記第三のペアの非同期リモートコピーの準備を開始し、前記第二のグループを構成する複数の第二論理ボリュームと、前記第三のグループを構成する複数の第三論理ボリュームとが同数であるかを判定し、判定の結果、同数である場合に、前記第三のペアの非同期リモートコピーが可能であると判断し、

前記第三のペアの非同期リモートコピーによる転送を開始する指示を受け付けるまで、前記転送を開始しない待機状態を維持することを特徴とする計算機システム。

10

【請求項 2】

請求項 1 記載の計算機システムであって、

前記第二の記憶システムは、前記第三のペアの非同期リモートコピーによる転送を開始する指示を受け付けると、前記予約された前記第三のペアの非同期リモートコピーを開始して、前記差分データを転送することを特徴とする計算機システム。

【請求項 3】

請求項 1 または請求項 2 記載の計算機システムであって、

前記ホストコンピュータは、グラフィカルユーザインターフェースと、前記第一のペア及び前記第二のペアの状態情報を有し、

前記グラフィカルユーザインターフェースに前記第一のペア状態と前記第二のペア状態とを提示して、ユーザからの入力を受け付けて、前記第三のペアの非同期リモートコピーの予約指示を行う、

20

ことを特徴とする計算機システム。

【請求項 4】

請求項 1 ないし請求項 3 のいずれか 1 つに記載の計算機システムであって、

非同期リモートコピー方式を用いた前記第一のグループのデータを前記第三のグループを構成する前記複数の第三論理ボリュームへ格納する処理として、前記第一の記憶システムは、前記ホストコンピュータから送信されたライトデータを受信し、ライトデータに対応したジャーナルを生成し、前記第三の記憶システムへ送信し、

非同期リモートコピー方式を用いた前記第一のグループのデータを前記第三のグループを構成する前記複数の第三論理ボリュームへ格納する処理として、前記第三の記憶システムは、前記ジャーナルを受信し、前記複数の第三論理ボリュームへ格納する、

30

ことを特徴とする計算機システム。

【請求項 5】

ホストコンピュータと、

第一のグループを構成する複数の第一論理ボリュームを有し、前記ホストコンピュータに接続する第一の記憶システムと、

前記複数の第一論理ボリュームに対応した複数の第二論理ボリュームを有する第二の記憶システムと、

前記第一のグループに対応する第三のグループを構成する複数の第三論理ボリュームを有する第三の記憶システムと、

40

から構成される計算機システムのデータ複製管理方法であって、

前記第二の記憶システムは、同期リモートコピー方式を用いて、前記複数の第一論理ボリュームのデータの複製を前記複数の第二論理ボリュームに格納し、

前記第三の記憶システムは、非同期リモートコピー方式を用いて、前記第一のグループのデータの複製を前記第三のグループを構成する前記複数の第三論理ボリュームに格納し、

前記第二の記憶システムは、同期リモートコピーを行う前記複数の第一論理ボリュームと前記複数の第二論理ボリュームとの第一のペアを管理し、

前記第三の記憶システムは、非同期リモートコピーを行う前記第一のグループと前記第

50

三のグループとの第二のペアを管理し、

前記第二の記憶システムは、

さらに、前記複数の第二論理ボリュームを含む第二のグループと、前記第三のグループとの第三のペアを管理し、

前記第一及び第二のペアにおけるリモートコピーが行われている間に、前記ホストコンピュータより、前記第二のグループのデータのうち前記第三のグループが格納していないデータである差分データを前記第三のグループに転送する、非同期リモートコピーの予約指示を受信すると、前記第三のペアの非同期リモートコピーの準備を開始し、前記第二のグループを構成する複数の第二論理ボリュームと、前記第三のグループを構成する複数の第三論理ボリュームとが同数であるかを判定し、判定の結果、同数である場合に、前記第三のペアの非同期リモートコピーが可能であると判断し、

前記第三のペアの非同期リモートコピーによる転送を開始する指示を受け付けるまで、前記転送を開始しない待機状態を維持する、

ことを特徴とする計算機システムのデータ複製管理方法。

【請求項 6】

請求項 5 記載の計算機システムのデータ複製管理方法であって、

前記第二の記憶システムは、前記第三のペアの非同期リモートコピーによる転送を開始する指示を受け付けると、前記予約された前記第三のペアの非同期リモートコピーを開始して、前記差分データを転送することを特徴とする計算機システムのデータ複製管理方法

。

【請求項 7】

請求項 5 または請求項 6 記載のデータ複製管理方法であって、

前記ホストコンピュータは、グラフィカルユーザインターフェースを有し、

前記ホストコンピュータは、グラフィカルユーザインターフェースと、前記第一のペア及び前記第二のペアの状態情報を有し、

前記グラフィカルユーザインターフェースに前記第一のペア状態と前記第二のペア状態とを提示して、ユーザからの入力を受け付けて、前記第三のペアの非同期リモートコピーの予約指示を行う、

ことを特徴とするデータ複製管理方法。

【請求項 8】

請求項 4 ないし請求項 7 のいずれか 1 つに記載のデータ複製管理方法であって、

前記非同期リモートコピー方式を用いた前記第一のグループのデータを前記第三のグループを構成する前記複数の第三論理ボリュームへ格納する処理として、前記第一の記憶システムは、前記ホストコンピュータから送信されたライトデータを受信し、ライトデータに対応したジャーナルを生成し、前記第三の記憶システムへ送信し、

前記非同期リモートコピー方式を用いた前記第一のグループのデータを前記第三のグループを構成する前記複数の第三論理ボリュームへ格納する処理として、前記第三の記憶システムは、前記ジャーナルを受信し、前記複数の第三論理ボリュームへ格納する、

ことを特徴とするデータ複製管理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、計算機システムにおける構成管理あるいは制御処理に関し、特に、リモートコピーを利用した複数のストレージシステムを含む構成管理あるいは制御処理に関する。

【背景技術】

【0002】

データストレージ市場では、大量のデータを格納した記憶システムが災害等で破壊されてもデータを喪失しない、いわゆるディザスタリカバリシステムが要求されている。このような市場の要求に応えるべく、リモートコピー技術を利用してデータをバックアップする計算機システムが提供されている。これは、十分に離れた二つの地点に設置された記憶

10

20

30

40

50

システムに同一のデータを格納するものである。一方の記憶システムのデータが更新されると、その更新は、リモートコピーによってもう一方の記憶システムに反映される。このため、二つの記憶システムのデータの同一性が確保される。

【 0 0 0 3 】

さらにデータの安全性を高めるため、相互に十分に離れた三つの地点に記憶システムを設置する計算機システムが特許文献 1 に開示されている。この計算機システムでは、通常の業務に使用される第 1 の記憶システムと、遠隔地の第 2 の記憶システムとの間のデータの同一性は、同期リモートコピーによって確保される。一方、第 1 の記憶システムと、遠隔地の第 3 の記憶システムとの間のデータの同一性は、非同期リモートコピーによって確保される。

10

【 0 0 0 4 】

災害等に起因する障害によって第 1 の記憶システムを業務に使用することができなくなった場合、第 2 の記憶システムが第 1 の記憶システムの業務を引き継ぐ。このとき、第 2 の記憶システムも使用することができない場合は、第 3 の記憶システムが第 1 の記憶システムの業務を引き継ぐ。

【 0 0 0 5 】

第 1 の記憶システムの業務を引き継いだ第 2 の記憶システムの運用が開始される前に、特許文献 1 では、第 2 の記憶システムと第 3 の記憶システムとの間のデータの同一性を確保する。第 2 の記憶システムの運用が開始された後は、第 2 の記憶システムのデータの更新をリモートコピーによって第 3 の記憶システムに反映させる。その結果、第 2 の記憶システムに障害が発生したときに、第 3 の記憶システムが第 2 の記憶システムの業務を引き継ぐことができる。

20

【 0 0 0 6 】

特許文献 2 記載のデータの更新方法では、この第 2 の記憶システムと第 3 の記憶システムとのデータの同一性の確保を行う時間を短縮する手段が開示されている。この技術は、第 2 の記憶システムと第 3 の記憶システムの同一性を確保する際に、相互の差異のデータをもう片方に反映することで、データの複製容量を削減し、結果として時間短縮を実現するものである。本明細書では、この特許文献 2 に示されるようなデータの同一性の確保技術を「デルタリシンク」と呼ぶ。

【 0 0 0 7 】

【特許文献 1】特開 2 0 0 3 - 1 2 2 5 0 9 号公報

【特許文献 2】特開 2 0 0 5 - 8 4 9 5 3 号公報

【発明の開示】

【発明が解決しようとする課題】

【 0 0 0 8 】

しかし、このようなデルタリシンクを行うには、特許文献 2 に示されるように、ジャーナル論理ボリュームの確保や、そのジャーナル論理ボリュームへの変更情報の格納、デルタリシンク時の変更情報の比較といったチェック処理など、従来のリモートコピーとは異なる様々な処理が必要になる。

【 0 0 0 9 】

更に、この第 1 の記憶システムが被災した後のチェック処理においてデルタリシンク不可との判定が出た場合、全てのデータが複製されることになり、特許文献 2 で解決を図った課題が再び発生することになっていた。即ち長時間のデータ複製により、第 2 の記憶システムを業務に使用することができなくなる。

40

【課題を解決するための手段】

【 0 0 1 0 】

上述の少なくとも一の課題を解決するために、本発明の一態様では、第 1 のホストコンピュータと通信可能に接続された第 1 の記憶システムと、第 2 のホストコンピュータ及び第 1 の記憶システムと通信可能に接続された第 2 の記憶システムと、第 1 及び第 2 の記憶システムと通信可能に接続された第 3 の記憶システムと、を備え、第 1 のホストコンピュ

50

ータから第1の記憶システムに書き込まれたデータは、第1の記憶システムから第2の記憶システムと第3の記憶システムにそれぞれ転送される計算機システムにおいて、第2の記憶システムは、第3の記憶システムにデータを転送するための記憶領域を有する転送設定を予め設定しておくことを特徴とする。

【0011】

さらに、別の態様として、前記第2の記憶システムは、前記第2の記憶システムと第3の記憶システム間のデータ転送の開始より前に、記憶領域及びデータ転送線及び転送設定情報のチェックを行い、その結果を前記転送設定に付随した情報としてホストコンピュータに報告することを特徴とする。

【発明の効果】

10

【0012】

本発明の上述の態様によれば、第2の記憶システムと第3の記憶システムにおけるデータ転送設定を予めホストコンピュータから監視することができ、デルタリシンクできない状態を解消することができる。

【発明を実施するための最良の形態】

【0013】

実施形態を図面により詳細に説明する。

【0014】

図1は、本発明の一実施形態の論理的な構成を示すブロック図である。本発明の一実施形態は、ホストコンピュータ180と記憶システム100Aを接続バス190により接続し、記憶システム100Aと記憶システム100Aに保存されたデータの複製を保持する記憶システム100Bと記憶システム100Cを接続バス200により接続した構成である。さらに、記憶システム100Bと記憶システム100Cは、接続バス200により接続されている。以下の説明において、複製対象のデータを保持する記憶システム100と複製データを保持する記憶システム100との区別を容易とするために、複製対象のデータを保持する記憶システム100を正記憶システム100A、複製データを保持する記憶システム100を副記憶システム100Bおよび副記憶システム100Cとよぶことがある。なお、記憶システムの記憶領域は、分割して管理されており、分割された記憶領域を論理ボリュームと呼ぶこととする。

20

【0015】

30

なお、本実施の形態において、記憶システム100は特許文献2記載のデータ更新方法を実装しているものとする。すなわち、記憶システム100は、記憶システム100A又はホストコンピュータ180に障害が発生し、記憶システムCを用いて業務を再開する場合に、記憶システムBと記憶システムCの間をジャーナルの差分のみの転送により複製対象と複製データの対とする「デルタリシンク」を実現する処理プログラムを実装している。

【0016】

論理ボリューム230の容量および記憶システム100内の物理的な格納位置（物理アドレス）は、記憶システム100に接続したコンピュータ等の保守端末もしくはホストコンピュータ180を用いて指定できる。各論理ボリューム230の物理アドレスは、後述するボリューム情報400に保存する。物理アドレスは、例えば、記憶システム100内の記憶装置150を識別する番号（記憶装置番号）と記憶装置内の記憶領域を一意に示す数値、例えば、記憶装置の記憶領域の先頭からの位置である。以下の説明では、物理アドレスは、記憶装置番号と記憶装置の記憶領域の先頭からの位置の組とする。以下の説明では、論理ボリュームは、1つの記憶装置の記憶領域であるが、論理アドレスと物理アドレスの変換により、1つの論理ボリュームを複数の記憶装置の記憶領域に対応づけることも可能である。

40

【0017】

記憶システム100が保存しているデータの参照、更新は、論理ボリュームを識別する番号（論理ボリューム番号）と記憶領域を一意に示す数値、例えば、論理ボリュームの記

50

憶領域の先頭からの位置により一意に指定することができ、以下、論理ボリューム番号と論理ボリュームの記憶領域の先頭からの位置（論理アドレス内位置）の組を論理アドレスとよぶ。

【 0 0 1 8 】

以下の説明において、複製対象のデータと複製データとの区別を容易とするために、複製対象の論理ボリューム 2 3 0 を正論理ボリューム、複製データである論理ボリューム 2 3 0 を副論理ボリュームとよぶこととする。一对の正論理ボリュームと副論理ボリュームをペアとよぶ。正論理ボリュームと副論理ボリュームの関係および状態等は後述するペア情報 5 0 0 に保存する。

【 0 0 1 9 】

論理ボリューム間のデータの更新順序を守るために、グループという管理単位を設ける。例えば、ホストコンピュータ 1 8 0 が、正論理ボリューム 1 のデータ 1 を更新し、その後、データ 1 を読み出し、データ 1 の数値を用いて、正論理ボリューム 2 のデータ 2 を更新する処理を行うとする。正論理ボリューム 1 から副論理ボリューム 1 へのデータ複製処理と、正論理ボリューム 2 から副論理ボリューム 2 へのデータ複製処理とが独立に行われる場合、副論理ボリューム 1 へのデータ 1 の複製処理より前に、副論理ボリューム 2 へのデータ 2 の複製処理が行われる場合がある。副論理ボリューム 2 へのデータ 2 の複製処理と副論理ボリューム 1 へのデータ 1 の複製処理との間に、故障等により副論理ボリューム 1 へのデータ 1 の複製処理が停止した場合、副論理ボリューム 1 と副論理ボリューム 2 のデータの整合性がなくなる。このような場合にも副論理ボリューム 1 と副論理ボリューム 2 のデータの整合性を保つために、データの更新順序を守る必要のある論理ボリュームは、同じグループに登録し、データの更新毎に、後述するグループ情報 6 0 0 の更新番号を割り当て、更新番号順に、副論理ボリュームに複製処理を行う。更新番号の代わりに更新時刻を用いてもよい。例えば、図 1 では、正記憶システム 1 0 0 A の論理ボリューム（DATA 1）と論理ボリューム（DATA 2）がグループ 1 を構成する。さらに、論理ボリューム（DATA 1）の複製である論理ボリューム（data 1）と論理ボリューム（DATA 2）の複製である論理ボリューム（data 2）は、副記憶システム 1 0 0 C 内でグループ 1 を構成する。同様に、論理ボリューム（DATA 1）の複製である論理ボリューム（COPY 1）と論理ボリューム（DATA 2）の複製である論理ボリューム（COPY 2）は、副記憶システム 1 0 0 B 内でグループ 1 を構成する。

【 0 0 2 0 】

データ複製対象である正論理ボリューム（DATA 1、DATA 2）のデータを更新する場合、副論理ボリューム（COPY 1、COPY 2）のデータを更新するために、正記憶システム 1 0 0 A は、後述するジャーナルを作成し、正記憶システム 1 0 0 A 内の論理ボリュームに保存する。本実施例の説明では、グループ毎にジャーナルのみを保存する論理ボリューム（以下、ジャーナル論理ボリュームとよぶ）を割り当てる。図 1 では、グループ 1 のジャーナル論理ボリュームは、論理ボリューム（JNL 1）である。

【 0 0 2 1 】

同様に、副記憶システム 1 0 0 C の副論理ボリューム（data 1、data 2）のデータを更新する場合、副記憶システム 1 0 0 C は、後述するジャーナルを作成し、副記憶システム 1 0 0 C 内のジャーナル論理ボリュームに保存する。図 1 では、グループ 1 のジャーナル論理ボリュームは、論理ボリューム（Jnl 1）である。

【 0 0 2 2 】

副記憶システム 1 0 0 B のグループにもジャーナル論理ボリュームを割り当てる。ジャーナル論理ボリュームは、正記憶システム 1 0 0 A から副記憶システム 1 0 0 B に転送したジャーナルを保存するために使用する。ジャーナルをジャーナル論理ボリュームに保存することにより、例えば、副記憶システム 1 0 0 B の負荷が高い場合、ジャーナル受信時に副論理ボリューム（COPY 1、COPY 2）のデータ更新を行わず、暫く後、副記憶システム 1 0 0 B の負荷が低い時に、副論理ボリューム（COPY 1、COPY 2）のデータを更新することもできる。さらに、接続線 2 0 0 が複数ある場合、正記憶システム 1

10

20

30

40

50

00Aから副記憶システム100Bへのジャーナルの転送を多重に行い、接続線200の転送能力を有効に利用することができる。更新順番のため、副記憶システム100Bに多くのジャーナルが溜まる可能性があるが、副論理ボリュームのデータ更新に直ぐに使用できないジャーナルは、ジャーナル論理ボリュームに格納することにより問題とならない。図1では、グループ1のジャーナル論理ボリュームは、論理ボリューム(JNL2)である。

#### 【0023】

ジャーナルは、ライトデータと更新情報とから構成する。更新情報は、ライトデータを管理するための情報で、ライト命令を受信した時刻(更新時刻)、グループ番号、後述するグループ情報600の更新番号、ライト命令の論理アドレス、ライトデータのデータサイズ、ライトデータを格納したジャーナル論理ボリュームの論理アドレス等からなる。更新情報は、ライト命令を受信した時刻と更新番号のどちらか一方のみを保持してもよい。ホストコンピュータ180からのライト命令内にライト命令の作成時刻が存在する場合は、ライト命令を受信した時刻の代わりに、当該ライト命令内の作成時刻を使用してもよい。図3と図21を用いて、ジャーナルの更新情報の例を説明する。更新情報310は、1999年3月17日の22時20分10秒に受信したライト命令を記憶する。当該ライト命令は、ライトデータを論理ボリューム番号1の記憶領域の先頭から700の位置に格納する命令であり、データサイズは300である。ジャーナルのライトデータは、論理ボリューム番号4(ジャーナル論理ボリューム)の記憶領域の先頭から1500の位置から格納されている。論理ボリューム番号1の論理ボリュームはグループ1に属し、グループ1のデータ複製開始から4番目のデータ更新であることがわかる。

#### 【0024】

ジャーナル論理ボリュームは、例えば、図3に示すように、更新情報を格納する記憶領域(更新情報領域)とライトデータを格納する記憶領域(ライトデータ領域)に分割して使用する。更新情報領域は、更新情報領域の先頭から、更新番号の順に格納し、更新情報領域の終端に達すると、更新情報領域の先頭から格納する。ライトデータ領域は、ライトデータ領域の先頭からライトデータを格納し、ライトデータ領域の終端に達すると、ライトデータ領域の先頭から格納する。更新情報領域およびライトデータ領域の比は、固定値でもよいし、保守端末あるいはホストコンピュータ180により設定可能としてもよい。これらの情報は、後述するポイント情報700内に保持する。以下の説明では、ジャーナル論理ボリュームを更新情報とライトデータの領域に分割して、ジャーナル論理ボリュームを使用するが、論理ボリュームの先頭から、ジャーナル、つまり、更新情報とライトデータを連続に格納する方式を採用してもよい。

#### 【0025】

図1を用いて、正記憶システム100Aの正論理ボリューム(DATA1)へのデータ更新を副記憶システム100Cの副論理ボリューム(data1)および副記憶システム100Bの副論理ボリューム(COPY1)に反映する動作について概説する。

#### 【0026】

(1) 正記憶システム100Aは、ホストコンピュータ180から正論理ボリューム(DATA1)内のデータに対するライト命令を受信すると、後述する命令受信処理210およびリードライト処理220によって、正論理ボリューム(DATA1)内のデータ更新と、ジャーナル論理ボリューム(JNL1)へのジャーナルの保存と、副記憶システム100Cに副記憶システム100C内の副論理ボリューム(data1)の対応するデータの更新(リモートライト命令)の命令を行う(図1の270)。

#### 【0027】

(2) 副記憶システム100Cは、正記憶システム100Aからのリモートライト命令を受け、後述する命令受信処理210およびリードライト処理220によって、副論理ボリューム(data1)内の対応するデータの更新と、ジャーナル論理ボリューム(JNL1)にジャーナルの保存を行う(図1の270)。

#### 【0028】

10

20

30

40

50

(3) 正記憶システム100Aは、リモートライト命令の応答を受けた後、ホストコンピュータにライト命令の終了を報告する。従って、正記憶システム100A内の正論理ボリューム(DATA1)と副記憶システム100C内の副論理ボリューム(data1)のデータは完全に一致する。このようなデータ複製を同期のデータ複製とよぶ。

【0029】

(4) 副記憶システム100Bは、後述するジャーナルリード処理240によって、正記憶システム100Aからジャーナルをリードし、リードライト処理220によって、ジャーナル論理ボリューム(JNL2)にジャーナルを保存する(図1の280)。

【0030】

(5) 正記憶システム100Aは、副記憶システム100Bからジャーナルリード命令を受信すると、後述する命令受信処理210およびリードライト処理220によって、ジャーナル論理ボリューム(JNL1)からジャーナルを読み出し、副記憶システム100Bに送信する(図1の280)。

【0031】

(6) 副記憶システム100Bは、後述するリストア処理250およびリードライト処理220によって、ポインタ情報700を用いて、更新番号の昇順に、ジャーナル論理ボリューム(JNL2)からジャーナルを読み出し、副論理ボリューム(COPY1)のデータを更新する(図1の290)。従って、正記憶システム100A内の正論理ボリューム(DATA1)と副記憶システム100B内の副論理ボリューム(COPY1)のデータは、正論理ボリューム(DATA1)の更新の暫く後には、完全に一致する。このようなデータ複製を非同期のデータ複製とよぶ。

【0032】

記憶システム100の内部構造を図2に示す。記憶システム100は、1つ以上のホストアダプタ110、1つ以上のディスクアダプタ120、1つ以上のキャッシュメモリ130、1つ以上の共有メモリ140、1つ以上の記憶装置150、1つ以上のコモンバス160、1つ以上の接続線170を備えて構成される。ホストアダプタ110、ディスクアダプタ120、キャッシュメモリ130、共有メモリ140はコモンバス160により相互間が接続されている。コモンバス160は、コモンバス160の障害時のために2重化されてもよい。ディスクアダプタ120と記憶装置150とは接続線170によって接続されている。また、図示していないが、記憶システム100の設定、監視、保守等を行うための保守端末が全てのホストアダプタ110とディスクアダプタ120とに専用線を用いて接続されている。

【0033】

ホストアダプタ110は、ホストコンピュータ180とキャッシュメモリ130間のデータ転送を制御する。ホストアダプタ110は、接続線190および接続線200によりホストコンピュータ180もしくは、他の記憶システム100と接続される。ディスクアダプタ120は、キャッシュメモリ130と記憶装置150との間のデータ転送を制御する。キャッシュメモリ130は、ホストコンピュータ180から受信したデータあるいは記憶装置150から読み出したデータを一時的に保持するメモリである。共有メモリ140は、記憶システム100内の全てのホストアダプタ110とディスクアダプタ120とが共有するメモリである。

【0034】

ボリューム情報400は、論理ボリュームを管理する情報であり、ボリューム状態、フォーマット形式、容量、同期ペア番号、非同期ペア番号、物理アドレスを保持する。図4にボリューム情報400の一例を示す。ボリューム情報400は、ホストアダプタ110およびディスクアダプタ120から参照可能なメモリ、例えば共有メモリ140に保存される。ボリューム状態は、“正常”、“正”、“副”、“異常”、“未使用”のいずれかを保持する。ボリューム状態が“正常”もしくは“正”である論理ボリューム230は、ホストコンピュータ180から正常にアクセス可能な論理ボリューム230であることを示す。ボリューム状態が“副”である論理ボリューム230は、ホストコンピュータ18

10

20

30

40

50

0からのアクセスを許可してもよい。ボリューム状態が“正”である論理ボリューム230は、データの複製が行われている論理ボリューム230であることを示す。ボリューム状態が“副”である論理ボリューム230は、複製に使用されている論理ボリューム230であることを示す。ボリューム状態が“異常”の論理ボリューム230は、障害により正常にアクセスできない論理ボリューム230であることを示す。障害とは、例えば、論理ボリューム230を保持する記憶装置150の故障である。ボリューム状態が“未使用”の論理ボリューム230は、使用していない論理ボリューム230であることを示す。同期ペア番号および非同期ペア番号は、ボリューム状態が“正”もしくは“副”の場合に有効であり、後述するペア情報500を特定するためのペア番号を保持する。保持するペア番号が存在しない場合は、無効値(例えば、“0”)を設定する。図4に示す例では、論理ボリューム1は、フォーマット形式がOPEN3、容量が3GB、記憶装置番号1の記憶装置150の記憶領域の先頭からデータが格納されており、アクセス可能であり、データの複製対象であることを示す。

10

#### 【0035】

図5にペア情報500の一例を示す。ペア情報500は、ペアを管理する情報であり、ペア状態、正記憶システム番号、正論理ボリューム番号、副記憶システム番号、副論理ボリューム番号、グループ番号、コピー済みアドレスを保持する。ペア情報500は、ホストアダプタ110およびディスクアダプタ120から参照可能なメモリ、例えば共有メモリ140に保存する。

#### 【0036】

20

ペア状態は、“正常”、“異常”、“未使用”、“コピー未”、“コピー中”、“デルタ待機”、“デルタ準備”、“デルタ異常”のいずれかを保持する。ペア状態が“正常”の場合は、正論理ボリューム230のデータ複製が正常に行われていることを示す。ペア状態が“異常”の場合は、障害により正論理ボリューム230の複製が行えないことを示す。障害とは、例えば、接続バス200の断線などである。ペア状態が“未使用”の場合は、当該ペア番号の情報は有効でないことを示す。ペア状態が“コピー中”の場合は、後述する初期コピー処理中であることを示す。ペア状態が“コピー未”の場合は、後述する初期コピー処理が未だ行われていないことを示す。ペア状態が“デルタ待機”の場合は、デルタリシンク可能な状態であることを示す。ペア状態が“デルタ準備”の場合は、デルタ待機に遷移する前の準備状態であることを示す。ペア状態が“デルタ異常”の場合は、デルタリシンクを行うためのジャーナル論理ボリューム等に障害が発生し、デルタリシンクや、デルタリシンクの準備が正常に行われていない状態を示す。正記憶システム番号は、正論理ボリューム230を保持する正記憶システム100Aを特定する番号を保持する。副記憶システム番号は、副論理ボリューム230を保持する副記憶システム100Bを特定する番号を保持する。グループ番号は、正記憶システムの場合は、正論理ボリュームが属するグループ番号を保持する。副記憶システムの場合は、副論理ボリュームが属するグループ番号を保持する。コピー済みアドレスは、後述する初期コピー処理にて説明する。図5のペア情報1は、データ複製対象が正記憶システムA内の正論理ボリューム1であり、データ複製先が副記憶システムB内の副論理ボリューム1であり、データ複製処理が行われていることを示す。

30

40

#### 【0037】

図6にグループ情報600の一例を示す。グループ情報600は、グループ状態、ペア集合、ジャーナル論理ボリューム番号、更新番号、複製種類、相手記憶システム番号、相手グループ番号、デルタ予約オプションを保持する。グループ情報600は、ホストアダプタ110およびディスクアダプタ120から参照可能なメモリ、例えば共有メモリ140に保存する。

#### 【0038】

グループ状態は、“正常”、“異常”、“未使用”、“停止”、“準備中”のいずれかを保持する。グループ状態が“正常”の場合は、ペア集合の少なくともひとつのペア状態が“正常”であることを示す。グループ状態が“異常”の場合は、ペア集合の全てのペア

50

状態が“異常”であることを示す。グループ状態が“未使用”の場合は、当該グループ番号の情報は有効でないことを示す。正記憶システムの場合、グループ状態の“停止”は、ジャーナルの作成を一時的に行わないことを示す。これは、グループ状態が“正常”の場合に、ジャーナルの作成を一時的に停止したい場合に使用する。副記憶システムの場合、グループ状態の“停止”は、ジャーナルリード処理を一時的に行わないことを示す。これは、グループ状態が“正常”の場合に、正記憶システムからジャーナルをリードすることを一時的に停止したい場合に使用する。グループ状態が“準備中”の場合は、後述するデータ複製開始処理中であることを示す。正記憶システムの場合、ペア集合はグループ番号が示すグループに属する全ての正論理ボリュームのペア番号を保持する。副記憶システムの場合、ペア集合はグループ番号が示すグループに属する全ての副論理ボリュームのペア番号を保持する。ジャーナル論理ボリューム番号は、当該グループ番号のグループに属するジャーナル論理ボリュームが存在しない場合は、無効値（例えば、“0”）を設定する。更新番号は、初期値は1であり、ジャーナルを作成した場合に変更する。更新番号は、ジャーナルの更新情報に記憶し、副記憶システム100Bにて、データの更新順を守るために使用する。

10

#### 【0039】

複製種類は、複製処理の種類を示す。本実施形態では、“同期”、“非同期”のいずれかを保持する。複製種類が“同期”の場合、副論理ボリュームの更新は、正論理ボリュームの更新処理のホストコンピュータへの書き込み完了報告前に行われる。従って、正論理ボリュームのデータと副論理ボリュームのデータはホストコンピュータからの書き込みの単位でみると完全完全に一致する。複製種類が“非同期”の場合、副論理ボリュームの更新は、正論理ボリュームの更新の後に行われ、必ずしもホストコンピュータへの書き込み完了報告前に行われるとは限らない。従って、正論理ボリュームのデータと副論理ボリュームのデータは、一致していない（副論理ボリュームのデータは正論理ボリュームの古いデータ）場合があり、暫く後、副論理ボリュームのデータは正論理ボリュームのデータと完全に一致する。上述した複製種類は、一例であり他の複製の種類でもよい。

20

#### 【0040】

正記憶システムの場合、相手記憶システム番号は、グループに属するペアの副論理ボリュームを保持する副記憶システム番号を保持する。副記憶システムの場合、相手記憶システム番号は、グループに属するペアの正論理ボリュームを保持する正記憶システム番号を保持する。正記憶システムの場合、相手グループ番号は、グループに属するペアの副論理ボリュームが属するグループ番号を保持する。副記憶システムの場合、相手グループ番号は、グループに属するペアの正論理ボリュームが属するグループ番号を保持する。

30

#### 【0041】

デルタ予約オプションは、当該グループがデルタリシンク用のグループとして登録されているかを示す情報を保持する。本実施の形態では、デルタ予約オプションが“0”の場合は通常のグループとして登録されており、“1”の場合はデルタリシンク用のグループとして登録されていることを表すものとする。例えば、図6のグループ情報1は、ペア情報1, 2から、正論理ボリューム1, 2と、ジャーナル論理ボリューム4から構成されており、通常のグループとして、正常にデータの複製処理（非同期）が行われていることを示す。

40

#### 【0042】

図7および図8にポインタ情報700の一例を示す。ポインタ情報700は、グループ毎に保持し、当該グループのジャーナル論理ボリュームを管理する情報であり、更新情報領域先頭アドレス、ライトデータ領域先頭アドレス、更新情報最新アドレス、更新情報最古アドレス、ライトデータ最新アドレス、ライトデータ最古アドレス、リード開始アドレス、リトライ開始アドレスを保持する。更新情報領域先頭アドレスは、ジャーナル論理ボリュームの更新情報を格納する記憶領域（更新情報領域）の先頭の論理アドレスを保持する。ライトデータ領域先頭アドレスは、ジャーナル論理ボリュームのライトデータを格納

50

する記憶領域（ライトデータ領域）の先頭の論理アドレスを保持する。更新情報最新アドレスは、次にジャーナルを格納する場合に、更新情報の保存に使用する先頭の論理アドレスを保持する。更新情報最古アドレスは、最古の（更新番号が小さい）ジャーナルの更新情報を保存する先頭の論理アドレスを保持する。ライトデータ最新アドレスは、次にジャーナルを格納する場合に、ライトデータの保存に使用する先頭の論理アドレスを保持する。ライトデータ最古アドレスは、最古の（更新番号が小さい）ジャーナルのライトデータを保存する先頭の論理アドレスを保持する。リード開始アドレスとリトライ開始アドレスは、正記憶システム 100A のみで使用し、後述するジャーナルリード受信処理にて使用する。

#### 【0043】

10

図7および図8のポインタ情報700の例では、ジャーナルの更新情報を保存する領域（更新情報領域）は、論理ボリューム4の記憶領域の先頭から699の位置までであり、ジャーナルのライトデータを保存する領域（ライトデータ領域）は、論理ボリューム4の記憶領域の700の位置から2699の位置までである。ジャーナルの更新情報は、論理ボリューム4の記憶領域の200の位置から499の位置まで保存されており、次のジャーナルの更新情報は、論理ボリューム4の記憶領域の500の位置から保存する。ジャーナルのライトデータは論理ボリューム4の記憶領域の1300の位置から2199の位置まで保存されており、次のジャーナルのライトデータは、論理ボリューム4の記憶領域の2200の位置から保存する。

#### 【0044】

20

下記の説明では、1つのグループに1つのジャーナル論理ボリュームを割り当てる形態で説明しているが、1つのグループに複数のジャーナル論理ボリュームを割り当ててもよい。例えば、1つのグループに2つのジャーナル論理ボリュームを割り当て、ジャーナル論理ボリューム毎にポインタ情報700を設け、交互にジャーナルを格納する。これにより、ジャーナルの記憶装置150への書き込みが分散でき、性能の向上が見込める。さらに、ジャーナルのリード性能も向上する。別の例としては、1つのグループに2つのジャーナル論理ボリュームを割り当て、通常は、1つのジャーナル論理ボリュームのみを使用する。もう一方のジャーナル論理ボリュームは、使用しているジャーナル論理ボリュームの性能が低下した場合もしくは故障して使用できない場合に使用する。性能が低下する例は、ジャーナル論理ボリュームが、複数の記憶装置150から構成され、RAID方式でデータを保持しており、RAIDを構成する記憶装置150の一台が故障中の場合である。

30

#### 【0045】

なお、上述のボリューム情報400、ペア情報500、グループ情報600、及びポインタ情報700等は、共有メモリ140に格納されていることが好ましい。しかし、本実施例は、この場合に限られず、これらの情報を、キャッシュメモリ130、ホストアダプタ110、ディスクアダプタ120、その他記憶装置150に集中して格納または分散して格納することもよい。

#### 【0046】

次に、正記憶システム100Aから副記憶システム100Bおよび副記憶システム100Cに対して、データ複製を開始する手順（データ複製開始処理）を図9、図10を用いて説明する。

40

#### 【0047】

（1）グループ作成について説明する（ステップ900）。ホストコンピュータ180は、正記憶システム100Aのグループ情報600を参照し、グループ状態が“未使用”のグループ番号Aを取得する。同様に、ホストコンピュータ180は、副記憶システム100B（もしくは副記憶システム100C）のグループ番号Bを取得する。ホストコンピュータ180は、正記憶システム100Aに、グループ作成指示を行う。グループ作成指示は、指示対象グループ番号A、相手記憶システム番号B、相手グループ番号B、複製種類、デルタ予約オプションからなる。デルタ予約オプションとは、当該グループを通常の

50

グループとして登録するか、デルタリシンク用のグループとして登録するかを指示するオプションである。本実施の形態では、デルタ予約オプションが“ 0 ”もしくは省略された場合は通常のグループとして登録し、“ 1 ”が記載された場合はデルタリシンク用のグループとして登録することを表すものとして説明を行う。このステップでの指示では、デルタ予約オプションは“ 0 ”である。

**【 0 0 4 8 】**

グループ作成指示を受けて、正記憶システム 1 0 0 A は、グループ情報 6 0 0 を変更する。具体的には、正記憶システム 1 0 0 A は、指示対象グループ番号 A のグループ情報 6 0 0 のグループ状態を“ 準備中 ”に、相手記憶システム番号を、指示された相手記憶システム番号 B に、相手グループ番号を、指示された相手グループ番号 B に、複製種類を指示された複製種類に、デルタ予約オプションは指示されたデルタ予約オプションに変更する。グループ情報 6 0 0 の更新番号を 1 ( 初期値 ) に設定する。さらに、正記憶システム 1 0 0 A は、相手記憶システム番号 B の記憶システムに対して、グループ作成指示を行う。グループ作成指示の指示対象グループ番号は相手グループ番号 B、相手記憶システム番号は正記憶システム 1 0 0 A の記憶システム番号、相手グループ番号は指示対象グループ番号 A、複製種類は指示された複製種類、デルタ予約オプションは、指示されたデルタ予約オプションとする。

10

**【 0 0 4 9 】**

( 2 ) ペア登録について説明する ( ステップ 9 1 0 ) 。ホストコンピュータ 1 8 0 は、データ複製対象を示す情報とデータ複製先を示す情報を指定し、ペア登録指示を正記憶システム 1 0 0 A に行う。データ複製対象を示す情報は、データ複製対象のグループ番号 A と正論理ボリューム番号 A である。データ複製先を示す情報は、複製データを保存する副記憶システム 1 0 0 B 内の副論理ボリューム番号 B である。

20

**【 0 0 5 0 】**

前記ペア登録指示を受けて、正記憶システム 1 0 0 A は、ペア情報 5 0 0 からペア情報が“ 未使用 ”のペア番号を取得し、ペア状態を“ コピー未 ”に、正記憶システム番号に正記憶システム 1 0 0 A を示す正記憶システム番号 A を、正論理ボリューム番号に指示された正論理ボリューム番号 A を、副記憶システム番号にグループ情報 6 0 0 のグループ番号 A の相手記憶システム番号を、副論理ボリューム番号に指示された副論理ボリューム番号 B を、グループ番号に指示されたグループ番号 A を設定する。指示されたグループ番号 A のグループ情報 6 0 0 のペア集合に前記取得したペア番号を追加し、正論理ボリューム番号 A のボリューム状態を“ 正 ”に変更する。

30

**【 0 0 5 1 】**

正記憶システム 1 0 0 A は、指示されたグループ番号 A のグループ情報 6 0 0 の相手記憶システムに対して、正記憶システム 1 0 0 A を示す正記憶システム番号 A、グループ番号 A のグループ情報 6 0 0 の相手グループ番号 B、正論理ボリューム番号 A、および副論理ボリューム番号 B を通知し、ペア登録を命令する。副記憶システム 1 0 0 B は、ペア情報 5 0 0 から未使用のペア番号を取得し、ペア状態を“ コピー未 ”に、正記憶システム番号に通知された正記憶システム番号 A を、正論理ボリューム番号に通知された正論理ボリューム番号 A を、副記憶システム番号に副記憶システム番号 B を、副論理ボリューム番号に通知された副論理ボリューム番号 B を、グループ番号に通知されたグループ番号 B を設定する。さらに、副記憶システム 1 0 0 B は、指示されたグループ番号 B のグループ情報のペア集合に前記取得したペア番号を追加し、副論理ボリューム番号 B のボリューム状態を“ 副 ”に変更する。

40

**【 0 0 5 2 】**

以上の動作を全てのデータ複製対象のペアに対して行う。

**【 0 0 5 3 】**

前記の説明では、論理ボリュームのグループへの登録と、論理ボリュームのペアの設定を同時に行う処理を説明したが、それぞれ個別に行ってもよい。

**【 0 0 5 4 】**

50

(3) ジャーナル論理ボリューム登録について説明する(ステップ920)。ホストコンピュータ180は、ジャーナルの保存に使用する論理ボリューム(ジャーナル論理ボリューム)をグループに登録する指示(ジャーナル論理ボリューム登録指示)を正記憶システム100Aに行う。ジャーナル論理ボリューム登録指示は、グループ番号と論理ボリューム番号からなる。

【0055】

正記憶システム100Aは、指示されたグループ番号のグループ情報600のジャーナル論理ボリューム番号に指示された論理ボリューム番号を登録する。さらに、当該論理ボリュームのボリューム情報400のボリューム状態を“正常”に設定する。

【0056】

同様に、ホストコンピュータ180は、副記憶システム100Bのボリューム情報400を参照し、副記憶システム100B、グループ番号B、ジャーナル論理ボリュームとして使用する論理ボリューム番号を指定し、ジャーナル論理ボリューム登録を正記憶システム100Aに行う。正記憶システム100Aは、ジャーナル論理ボリューム登録指示を副記憶システム100Bに転送する。副記憶システム100Bは、指示されたグループ番号Bのグループ情報600のジャーナル論理ボリューム番号に指示された論理ボリューム番号を登録する。さらに当該論理ボリュームのボリューム情報400のボリューム状態を“正常”に設定する。

【0057】

ホストコンピュータ180は、グループ番号、ジャーナル論理ボリュームとして使用する論理ボリューム番号を指定し、ジャーナル論理ボリューム登録指示を副記憶システム100Bに行ってもよい。副記憶システム100Cに対しても同様に行う。あるいは、記憶システムごとにホストコンピュータや保守端末が接続されており、それぞれのホストコンピュータや保守端末が、それぞれ接続されている記憶システムに指示を行ってもよい。

【0058】

以上の動作を全てのジャーナル論理ボリュームとして使用する論理ボリュームに対して行う。ステップ910とステップ920の順は逆でもよい。

【0059】

(4) データ複製処理開始について説明する(ステップ930)。ホストコンピュータ180は、データ複製処理を開始する複製種類が同期のグループ番号Cと複製種類が非同期のグループ番号Bを指定し、データ複製処理の開始を正記憶システム100Aに指示する。正記憶システム100Aは、グループBに属する全てのペア情報400のコピー済みアドレスを0に設定する。

【0060】

正記憶システム100Aは、グループ番号Bのグループ情報600の相手記憶システム100Bに対し、グループ番号Bのグループ情報600の相手グループ番号のグループ状態を“正常”に変更する事と後述するジャーナルリード処理およびリストア処理の開始を指示する。正記憶システム100Aは、グループ番号Cのグループ情報600の相手記憶システム100Cに対し、グループ番号Cのグループ情報600の相手グループ番号のグループ状態を“正常”に変更する事を指示する。

【0061】

正記憶システム100Aは、グループ番号Cのグループ状態とグループ番号Bのグループ状態を“正常”に変更し、後述する初期コピー処理を開始する。

【0062】

前記の説明では、同期のデータ複製処理と非同期のデータ複製処理の開始を同時に指示していたが、それぞれ個別に行ってもよい。

【0063】

(5) 初期コピー終了について説明する(ステップ940)。

【0064】

初期コピーが終了すると、正記憶システム100Aは、初期コピー処理の終了を副記憶

10

20

30

40

50

システム 1 0 0 B および副記憶システム 1 0 0 C に通知する。副記憶システム 1 0 0 B および副記憶システム 1 0 0 C は、グループ B およびグループ C に属する全ての副論理ボリュームのペア状態を“ 正常 ”に変更する。

【 0 0 6 5 】

( 6 ) デルタリシンク用のグループ作成について説明する ( ステップ 9 5 0 ) 。 ホストコンピュータ 1 8 0 は、デルタリシンク用のグループ作成指示を行う。グループ作成指示は、指示対象グループ番号 C 、相手記憶システム番号 B 、相手グループ番号 B 、複製種類、デルタ予約オプションからなる。ここでのデルタ予約オプションは“ 1 ”である。

【 0 0 6 6 】

グループ作成指示を受け取った副記憶システム 1 0 0 C は、グループ情報 6 0 0 を変更する。具体的には、副記憶システム 1 0 0 C は、指示対象グループ番号 C のグループ情報 6 0 0 のグループ状態を“ 停止 ”に、相手記憶システム番号を指示された相手記憶システム番号 B に、相手グループ番号を指示された相手グループ番号 B に、複製種類を指示された複製種類に、デルタ予約オプションは指示されたデルタ予約オプションに変更する。グループ情報 6 0 0 の更新番号を 1 ( 初期値 ) に設定する。さらに、副記憶システム 1 0 0 C は、相手記憶システム番号 B の記憶システムに対して、グループ作成指示を行う。グループ作成指示の指示対象グループ番号は相手グループ番号 B 、相手記憶システム番号は副記憶システム 1 0 0 C の記憶システム番号、相手グループ番号は指示対象グループ番号 C 、複製種類は指示された複製種類、デルタ予約オプションは、指示されたデルタ予約オプションとする。

( 7 ) デルタリシンク用のペア登録について説明する ( ステップ 9 6 0 ) 。 ホストコンピュータ 1 8 0 は、データ複製対象を示す情報とデータ複製先を示す情報を指定し、ペア登録指示を副記憶システム 1 0 0 C に行う。データ複製対象を示す情報は、データ複製対象のグループ番号 C と正論理ボリューム番号 C である。データ複製先を示す情報は、複製データを保存する副記憶システム 1 0 0 B 内の副論理ボリューム番号 B である。

【 0 0 6 7 】

前記ペア登録指示を受けて、副記憶システム 1 0 0 C は、ペア情報 5 0 0 からペア情報が“ 未使用 ”のペア番号を取得し、対応するグループ番号 C のデルタ予約オプションが“ 1 ”のためペア状態を“ デルタ準備 ”に、正記憶システム番号に副記憶システム 1 0 0 C を示す正記憶システム番号 C を、正論理ボリューム番号に指示された正論理ボリューム番号 C を、副記憶システム番号にグループ情報 6 0 0 のグループ番号 C の相手記憶システム番号を、副論理ボリューム番号に指示された副論理ボリューム番号 B を、グループ番号に指示されたグループ番号 C を設定する。指示されたグループ番号 C のグループ情報 6 0 0 のペア集合に前記取得したペア番号を追加する。

【 0 0 6 8 】

副記憶システム 1 0 0 C は、指示されたグループ番号 C のグループ情報 6 0 0 の相手記憶システムに対して、副記憶システム 1 0 0 C を示す正記憶システム番号 C 、グループ番号 C のグループ情報 6 0 0 の相手グループ番号 B 、正論理ボリューム番号 C 、および副論理ボリューム番号 B を通知し、ペア登録を命令する。副記憶システム 1 0 0 B は、ペア情報 5 0 0 から未使用のペア番号を取得し、ペア状態を対応するグループ番号 B のデルタ予約オプションが“ 1 ”のためペア状態を“ デルタ準備 ”に、正記憶システム番号に通知された正記憶システム番号 C を、正論理ボリューム番号に通知された正論理ボリューム番号 C を、副記憶システム番号に副記憶システム番号 B を、副論理ボリューム番号に通知された副論理ボリューム番号 B を、グループ番号に通知されたグループ番号 B を設定する。さらに、副記憶システム 1 0 0 B は、指示されたグループ番号 B のグループ情報のペア集合に前記取得したペア番号を追加する。

【 0 0 6 9 】

以上の動作を全てのデルタリシンク対象のペアに対して行う。

( 8 ) デルタ予約処理について説明する ( ステップ 9 7 0 ) 。 ホストコンピュータ 1 8 0 は、デルタ予約処理を開始する複製種類が非同期でデルタ予約オプションが“ 1 ”の

グループ番号Cを指定し、デルタ予約処理の開始を副記憶システム100Cに指示する。なお、デルタリシンクは、正記憶システム100Aの正論理ボリュームAとそれぞれペアになっている副記憶システム100Bの副論理ボリュームBと、副記憶システム100Cの副論理ボリュームCと、の差分だけをコピーしてデータの同期を取る処理である。

【0070】

指示を受け取った副記憶システム100Cは、デルタリシンクが可能か否かをチェックする。すなわち、まずデルタリシンク用に登録したペアを複製先としている2つのグループの複製対象が一致し、かつ、ペア数も一致しているかを、判定する。デルタリシンクが可能な状態と判定した場合は、当該グループ番号のペア情報500のペア状態を“デルタ待機”に変更する。以上がデルタ予約処理である。

10

【0071】

次に、ステップ970の判定処理の詳細について図46を用いて説明する。

【0072】

まず、記憶システム100Cは、デルタ予約指定のグループ情報に記載されたペア集合の、最初のペア番号を取得する(ステップ4610)。

【0073】

記憶システム100Cは、ペア情報500の中から、指定されたグループ番号Cであるペアの正記憶システム番号Cと正論理ボリューム番号Cと同一である副記憶システム番号と副論理ボリューム番号を持つペア情報を抽出し、そこから正記憶システム番号A1と正論理ボリューム番号A1を取得する(ステップ4630)。次に記憶システム100Cは、正記憶システム番号Cと正論理ボリューム番号Cに対応する副記憶システム番号Bである副記憶システム100Bから、正記憶システム番号Cと正論理ボリューム番号Cに対応する副記憶システム番号Bと副論理ボリュームBを複製先として有し、デルタ予約オプションが“1”でないグループ番号もつペアの正記憶システム番号A2と、正論理ボリューム番号A2を取得する(ステップ4640)。そして記憶システム100Cは、取得した正記憶システム番号A1及び正論理ボリューム番号A1と、正記憶システム番号A2と正論理ボリューム番号A2が一致するかを判定する(ステップ4650)。この判定を、ステップ4660、ステップ4620を経て指定されたグループ番号Cであるペア全てに対して行う。

20

【0074】

この判定(ステップ4650)で不一致がでた場合は、デルタ予約処理をエラーで完了する。

30

【0075】

指定されたペアの比較が終了したかを判定し(ステップ4670)、判定結果、終了していた場合、デルタリシンクを行う場合に必要なジャーナルデータの転送が行われているかを比較し判定する。

【0076】

まず、記憶システム100Cは、グループ情報600の中で、指定されたグループ番号Cでデルタ予約オプションが“1”であるグループ番号に記載されている相手記憶システム番号Bである副記憶システム100Bから、当該グループ番号に記載されている相手グループ番号Bであるグループに属するペアの副論理ボリュームと同一の副論理ボリュームであるペアを持つグループの更新番号Bに1を足した数値B1を取得する(ステップ4672)。次に、記憶システム100Cは、ポイント情報700を参照し、更新番号Bのジャーナルを保持しているかを調べる。記憶システム100Cは、ポイント情報の更新情報最古アドレスの更新情報を記憶装置150から読み出し、最古(最小)の更新番号Cを得る(ステップ4674)。記憶システムCは、ステップ4672で取得した値とステップ4674で取得した値とを比較し(ステップ4676)、更新番号Cが取得した更新番号B1と等しい、もしくは小さい場合は、記憶システム100Cが更新番号Bのジャーナルを保持しているので、デルタリシンクが可能な状態と判定する(ステップ4678)。

40

【0077】

50

図10は、初期コピー処理のフローチャートである。初期コピー処理は、データ複製対象の正論理ボリュームの全記憶領域に対し、ペア情報500のコピー済みアドレスを用い、記憶領域の先頭から順に、単位サイズ毎にジャーナルを作成する。コピー済みアドレスは、初期値は0であり、ジャーナルの作成毎に、作成したデータ量を加算する。論理ボリュームの記憶領域の先頭から、コピー済みアドレスの一つ前までは、初期コピー処理にてジャーナルは作成済みである。初期コピー処理を行うことにより、正論理ボリュームの更新されていないデータを副論理ボリュームに転送することが可能となる。以下の説明では、正記憶システム100A内のホストアダプタAが処理を行うように記載しているが、ディスクアダプタ120が行ってもよい。

【0078】

10

(1) 正記憶システム100A内のホストアダプタAは、処理対象の非同期複製のグループBに属するペアでペア状態が“コピー未”である正論理ボリュームAを得、ペアの状態を“コピー中”に変更し、以下の処理を繰り返す(ステップ1010、1020)。正論理ボリュームAが存在しない場合は、処理を終了する(ステップ1030)。

【0079】

(2) ステップ1020にて、論理ボリュームAが存在した場合、ホストアダプタAは、単位サイズ(例えば、1MB)のデータを対象にジャーナルを作成する。ジャーナル作成処理は後述する(ステップ1040)。

【0080】

(3) ホストアダプタAは、正論理ボリュームAの同期ペアである副論理ボリュームのデータを更新するために、同期ペアである副論理ボリュームを保持する副記憶システムCにリモートライト命令を送信する。リモートライト命令は、ライト命令、論理アドレス(論理ボリュームは同期ペア番号の副論理ボリュームC、論理ボリューム内位置はコピー済みアドレス)、データ量(単位サイズ)、ステップ1040で使用した更新番号を含んでいる。なお、更新番号の代わりにジャーナルを作成した時刻でもよい(ステップ1045)。リモートライト命令を受信した場合の副記憶システムCの動作は後述する命令受信処理210にて説明する。

20

【0081】

(4) ホストアダプタAは、リモートライト命令の応答を受信後、コピー済みアドレスに作成したジャーナルのデータサイズを加算する(ステップ1050)。

30

【0082】

(5) コピー済みアドレスが、正論理ボリュームAの容量に達するまで、上記処理を繰り返す(ステップ1060)。コピー済みアドレスが、正論理ボリュームAの容量と等しくなった場合、正論理ボリュームAの全記憶領域に対してジャーナルを作成したため、ペア状態を“正常”に更新し、他の正論理ボリュームの処理を開始する(ステップ1070)。

【0083】

前記のフローチャートでは、論理ボリュームを1つずつ対象とするように説明したが、複数の論理ボリュームを同時に処理してもよい。

【0084】

40

図11は命令受信処理210の処理を説明する図、図12は命令受信処理210のフローチャート、図13はジャーナル作成処理のフローチャート、図23はリモートライト命令受信処理のフローチャート、図24はジャーナル複製処理のフローチャートである。以下、これらを用いて、正記憶システム100Aが、ホストコンピュータ180からデータ複製対象の論理ボリューム230にライト命令を受信した場合の動作について説明する。

【0085】

(1) 記憶システム100A内のホストアダプタAは、ホストコンピュータからアクセス命令を受信する。アクセス命令は、リード、ライト、後述するジャーナルリード等の命令、命令対象の論理アドレス、データ量等を含んでいる。以下、アクセス命令内の論理アドレスを論理アドレスA、論理ボリューム番号を論理ボリュームA、論理ボリューム内位

50

置を論理ボリューム内位置 A、データ量をデータ量 A とする (ステップ 1 2 0 0)。

【 0 0 8 6 】

( 2 ) ホストアダプタ A は、アクセス命令を調べる (ステップ 1 2 1 0、1 2 1 5、1 2 2 8)。ステップ 1 2 1 5 の調べで、アクセス命令がジャーナルリード命令の場合は、後述するジャーナルリード受信処理を行う (ステップ 1 2 2 0)。アクセス命令が、リモートライト命令の場合は、後述するリモートライト命令受信処理を行う (ステップ 2 3 0 0)。それら以外の命令、例えば、リード命令の場合は、従来技術と同じようにリード処理を行い、リモートコピー制御命令であれば、指示に対応する処理を行う (ステップ 1 2 3 0)。

【 0 0 8 7 】

( 3 ) ステップ 1 2 1 0 の調べで、アクセス命令がライト命令の場合は、論理ボリューム A のボリューム情報 4 0 0 を参照し、ボリューム状態を調べる (ステップ 1 2 4 0)。ステップ 1 2 4 0 の調べで、論理ボリューム A のボリューム状態が、“正常”もしくは“正”以外の場合は、論理ボリューム A へのアクセスは不可能なため、ホストコンピュータ 1 8 0 に異常終了を報告する (ステップ 1 2 4 5)。

【 0 0 8 8 】

( 4 ) ステップ 1 2 4 0 の調べで、論理ボリューム A のボリューム状態が、“正常”、“正”のいずれかの場合は、ホストアダプタ A は、キャッシュメモリ 1 3 0 を確保し、ホストコンピュータ 1 8 0 にデータ受信の準備ができたことを通知する。ホストコンピュータ 1 8 0 は、その通知を受け、ライトデータを正記憶システム 1 0 0 A に送信する。ホストアダプタ A は、ライトデータを受信し、当該キャッシュメモリ 1 3 0 に保存する (ステップ 1 2 5 0、図 1 1 の 1 1 0 0)。

【 0 0 8 9 】

( 5 ) ホストアダプタ A は、論理ボリューム A のボリューム情報、ペア情報、グループ情報を参照し、論理ボリューム A が非同期複製対象かどうかを調べる (ステップ 1 2 6 0)。ステップ 1 2 6 0 の調べで、論理ボリューム A のボリューム状態が“正”かつ、論理ボリューム A の非同期ペア番号のペアのペア状態が“正常”かつ、当該ペアが属するグループのグループ状態が“正常”である場合は、論理ボリューム A が非同期複製対象であるため、後述するジャーナル作成処理を行う (ステップ 1 2 6 5)。

【 0 0 9 0 】

( 6 ) ホストアダプタ A は、論理ボリューム A のボリューム情報、ペア情報、グループ情報を参照し、論理ボリューム A が同期複製対象かどうかを調べる (ステップ 1 2 6 7)。ステップ 1 2 6 7 の調べで、論理ボリューム A のボリューム状態が“正”かつ、論理ボリューム A の同期ペア番号のペアのペア状態が“正常”かつ、当該ペアが属するグループのグループ状態が“正常”である場合は、論理ボリューム A が同期複製対象であるため、同期ペア番号のペアの副記憶システム C に、ホストコンピュータ 1 8 0 から受信したライトデータを保存するリモートライト命令を送信する (図 1 1 の 1 1 8 5)。リモートライト命令は、ライト命令、論理アドレス (論理ボリューム番号は、同期ペア番号のペアの副論理ボリューム C、論理ボリューム内位置は論理ボリューム内位置 A)、データ量 A、ステップ 1 2 6 5 で使用した更新番号を含んでいる。なお、更新番号の代わりにホストコンピュータ 1 8 0 からライト命令を受信した時刻でもよい。ステップ 1 2 6 7 の調べで、同期複製対象の論理ボリュームでない場合、もしくは、ステップ 1 2 6 8 のジャーナル作成処理が成功しなかった場合は、更新番号に無効を示す数値“0”を設定する。

【 0 0 9 1 】

( 7 ) ステップ 1 2 6 7 もしくは、ステップ 1 2 6 8 のリモートライト命令の応答を受信した後、ホストアダプタ A は、ディスクアダプタ 1 2 0 にライトデータを論理アドレス A に対応する記憶装置 1 5 0 の記憶領域に書き込むことを命令し (図 1 1 の 1 1 6 0)、ホストコンピュータ 1 8 0 に終了報告する (ステップ 1 2 7 0、1 2 8 0)。その後、当該ディスクアダプタ 1 2 0 は、リードライト処理により、前記記憶領域にライトデータを保存する (図 1 1 の 1 1 7 0)。

10

20

30

40

50

## 【 0 0 9 2 】

次に、ジャーナル作成処理について説明する。

## 【 0 0 9 3 】

( 1 ) ホストアダプタ A は、ジャーナル論理ボリュームのボリューム状態を調べる ( ステップ 1 3 1 0 ) 。ステップ 1 3 1 0 の調べで、ジャーナル論理ボリュームのボリューム状態が、“異常”の場合は、ジャーナル論理ボリュームにジャーナルの格納が不可能なため、グループ状態を“異常”に変更し、処理を終了する ( ステップ 1 3 1 5 ) 。この場合、ジャーナル論理ボリュームを正常な論理ボリュームに変更する等を行う。

## 【 0 0 9 4 】

( 2 ) ステップ 1 3 1 0 の調べで、ジャーナル論理ボリュームが正常である場合、ジャーナル作成処理を継続する。ジャーナル作成処理は、初期コピー処理内の処理であるか、命令受信処理内の処理であるかによって処理が異なる ( ステップ 1 3 2 0 ) 。ジャーナル作成処理が命令受信処理内の処理の場合は、ステップ 1 3 3 0 からの処理を行う。ジャーナル作成処理が初期コピー処理内の場合は、ステップ 1 3 7 0 からの処理を行う。

## 【 0 0 9 5 】

( 3 ) ジャーナル作成処理が命令受信処理内の処理の場合、ホストアダプタ A は、ライト対象の論理アドレス A が、初期コピー処理の処理対象となったかを調べる ( ステップ 1 3 3 0 ) 。論理ボリューム A のペア状態が“コピー未”の場合は、後に初期コピー処理にてジャーナル作成処理が行われるため、ジャーナルを作成せずに処理を終了する ( ステップ 1 3 3 5 ) 。論理ボリューム A のペア状態が“コピー中”の場合は、コピー済みアドレスが論理アドレス内位置 A と等しいもしくは、小さい場合は、後に初期コピー処理にてジャーナル作成処理が行われるため、ジャーナルを作成せずに処理を終了する ( ステップ 1 3 3 5 ) 。上記以外、つまり、論理ボリューム A のペア状態が“コピー中”かつコピー済みアドレスが論理アドレス内位置 A より大きい場合もしくは、論理ボリューム A のペア状態が“正常”の場合は、既に初期コピー処理が完了しているため、ジャーナル作成処理を継続する。

## 【 0 0 9 6 】

( 4 ) 次に、ホストアダプタ A は、ジャーナルがジャーナル論理ボリュームに格納可能であるかを調べる。ポインタ情報 7 0 0 を用い、更新情報領域の未使用領域の有無を調べる ( ステップ 1 3 4 0 ) 。ポインタ情報 7 0 0 の更新情報最新アドレスと更新情報最古アドレスが等しい場合は、更新情報領域に未使用領域が存在しないため、ジャーナル作成失敗として処理を終了する ( ステップ 1 3 9 0 ) 。

## 【 0 0 9 7 】

ステップ 1 3 4 0 の調べで、更新情報領域に未使用領域が存在する場合は、ポインタ情報 7 0 0 を用い、ライトデータ領域にライトデータが格納できるかを調べる ( ステップ 1 3 4 5 ) 。ライトデータ最古アドレスが、ライトデータ最新アドレスとライトデータ最新アドレスにデータ量 A を足した数値の範囲にある場合、当該ライトデータをライトデータ領域に格納できないため、ジャーナル作成失敗として処理を終了する ( ステップ 1 3 9 0 ) 。

## 【 0 0 9 8 】

( 5 ) ジャーナルが格納可能である場合、ホストアダプタ A は、更新番号と更新情報を格納する論理アドレスとライトデータを格納する論理アドレスを取得し、更新情報をキャッシュメモリ 1 3 0 内に作成する。更新番号は、対象グループのグループ情報 6 0 0 から取得し、1 を足した数値をグループ情報 6 0 0 の更新番号に設定する。更新情報を格納する論理アドレスは、ポインタ情報 7 0 0 の更新情報最新アドレスであり、更新情報のサイズを足した数値をポインタ情報 7 0 0 の更新情報最新アドレスに設定する。ライトデータを格納する論理アドレスは、ポインタ情報 7 0 0 のライトデータ最新アドレスであり、ライトデータ最新アドレスにデータ量 A を足した数値をポインタ情報 7 0 0 のライトデータ最新アドレスに設定する。

## 【 0 0 9 9 】

ホストアダプタ A は、上記取得した数値とグループ番号、ライト命令を受信した時刻、ライト命令内の論理アドレス A、データ量 A を更新情報に設定する（ステップ 1350、図 11 の 1120）。例えば、図 6 に示すグループ情報 600、図 7 に示すポインタ情報 700 の状態で、グループ 1 に属する正論理ボリューム 1 の記憶領域の先頭から 800 の位置にデータサイズ 100 のライト命令を受信した場合、図 22 に示す更新情報を作成する。グループ情報の更新番号は 6、ポインタ情報の更新情報最新アドレスは 600（更新情報のサイズは 100 とする）、ライトデータ最新アドレスは 2300 となる。

【0100】

（6）ホストアダプタ A は、ディスクアダプタ 120 に、ジャーナルの更新情報とライトデータを記憶装置 150 に書き込むことを命令し、正常終了する（ステップ 1360、図 11 の 1130、1140、1150）。

10

【0101】

（7）ジャーナル作成処理が、初期コピー処理内の処理の場合は、ステップ 1370 からの処理を行う。ホストアダプタ A は、ジャーナルが作成可能であることを調べる。ポインタ情報 700 を用い、更新情報領域の未使用領域の有無を調べる（ステップ 1370）。ポインタ情報 700 の更新情報最新アドレスと更新情報最古アドレスが等しい場合は、更新情報領域に未使用領域が存在しないため、ジャーナル作成失敗として処理を終了する（ステップ 1390）。本実施例で示した初期コピー処理の場合、ジャーナルのライトデータは、正論理ボリュームからリードし、ライトデータ領域は使用しないため、ライトデータ領域の未使用領域の確認は不要である。

20

【0102】

（8）ステップ 1370 の調べで、ジャーナルが作成可能である場合、ホストアダプタ A は、更新情報をキャッシュメモリ 130 内に作成する。更新情報のライト命令を受信した時刻は、更新番号を取得した時刻を設定する。グループ番号は、論理ボリュームの非同期ペア番号が属するグループ番号を設定する。更新番号は、グループ情報 600 から取得し、1 を足した数値をグループ情報 600 の更新番号に設定する。ライト命令の論理アドレスとライトデータを格納したジャーナル論理ボリュームの論理アドレスは、初期コピー処理対象の論理アドレス（ペア情報のコピー済みアドレス）を設定する。ライトデータのデータサイズは、初期コピー処理の単位サイズを設定する。更新情報を格納する論理アドレスは、ポインタ情報 700 の更新情報最新アドレスの位置であり、更新情報のサイズを足した数値をポインタ情報 700 の更新情報最新アドレスに設定する（ステップ 1380、図 11 の 1120）。

30

【0103】

（9）ホストアダプタ A は、ディスクアダプタ 120 に、更新情報を記憶装置 150 に書き込むことを命令し、正常終了する（ステップ 1385、図 11 の 1140、1150）。

【0104】

上記説明では、更新情報をキャッシュメモリ 130 内に存在するように記載しているが、共有メモリ 140 内等に格納してもよい。

【0105】

40

ライトデータの記憶装置 150 への書き込みは、非同期、つまり、ステップ 1360 およびステップ 1385 の直後でなくともよい。ただし、ホストコンピュータ 180 が、論理アドレス A にライト命令を再び行った場合、ジャーナルのライトデータが上書きされるため、ホストコンピュータ 180 からライトデータを受信する前に、ジャーナルのライトデータは、更新情報のジャーナル論理ボリュームの論理アドレスに対応する記憶装置 150 に書き込む必要がある。もしくは、別のキャッシュメモリに退避し、後に更新情報のジャーナル論理ボリュームの論理アドレスに対応する記憶装置 150 に書き込みを行ってもよい。

【0106】

前述したジャーナル作成処理では、ジャーナルを記憶装置 150 に保存するとしていた

50

が、ジャーナル用に予め一定量のキャッシュメモリ 130 を用意しておき、当該キャッシュメモリを全て使用してから、記憶装置 150 にジャーナルを保存してもよい。ジャーナル用のキャッシュメモリ量は、例えば、保守端末から指定する。

【0107】

次に、副記憶システム 100C のホストアダプタ C が、正記憶システム 100A からリモートライト命令を受信した場合の処理（リモートライト命令受信処理）について説明する。リモートライト命令は、ライト命令、論理アドレス（副論理ボリューム C、論理ボリューム内位置 A）、データ量 A、更新番号を含んでいる。

【0108】

（1）副記憶システム 100C 内のホストアダプタ C は、論理ボリューム C のボリューム情報 400 を参照し、副論理ボリューム C のボリューム状態を調べる（ステップ 2310）。ステップ 2310 の調べで、論理ボリューム C のボリューム状態が、“副”以外の場合は、論理ボリューム C へのアクセスは不可能なため、正記憶システム 100A に異常終了を報告する（ステップ 2315）。 10

【0109】

（2）ステップ 2310 の調べで、論理ボリューム C のボリューム状態が、“副”の場合は、ホストアダプタ C は、キャッシュメモリ 130 を確保し、正記憶システム 100A にデータ受信の準備ができたことを通知する。正記憶システム 100A は、その通知を受け、ライトデータを副記憶システム 100C に送信する。ホストアダプタ C は、ライトデータを受信し、当該キャッシュメモリ 130 に保存する（ステップ 2320）。 20

【0110】

（3）リモートライト命令に含まれる更新番号を調べ、無効値“0”の場合は、正記憶システム 100A にてジャーナルが作成されていないため、ステップ 2400 のジャーナル複製処理を行わない（ステップ 2330）。

【0111】

（4）リモートライト命令に含まれる更新番号を調べ、有効値（“0”以外）の場合は、ジャーナル論理ボリュームのボリューム状態を調べる。ジャーナル論理ボリュームのボリューム状態が“異常”の場合は、ジャーナル論理ボリュームにジャーナルの格納が不可能なため、ステップ 2400 のジャーナル複製処理を行わない（ステップ 2340）。 30

【0112】

（5）ステップ 2340 の調べで、ジャーナル論理ボリュームのボリューム状態が“正常”の場合は、後述するジャーナル複製処理 2400 を行う。

【0113】

（6）ホストアダプタ C は、ディスクアダプタ 120 にライトデータをリモートライト命令の論理アドレスに対応する記憶装置 150 の記憶領域に書き込むことを命令し、正記憶システム A に終了報告する（ステップ 2360、2370）。その後、当該ディスクアダプタ 120 は、リードライト処理により、前記憶領域にライトデータを保存する。

【0114】

次に、ジャーナル複製処理 2400 について説明する。

【0115】

（1）ホストアダプタ C は、ジャーナルがジャーナル論理ボリュームに格納可能であることを調べる。ポインタ情報 700 を用い、更新情報領域の未使用領域の有無を調べる（ステップ 2410）。ポインタ情報 700 の更新情報最新アドレスと更新情報最古アドレスが等しい場合は、更新情報領域に未使用領域が存在しないため、最古のジャーナルの記憶領域を開放し、更新情報領域を確保する（ステップ 2415）。次に、ポインタ情報 700 を用い、ライトデータ領域にライトデータが格納できるかを調べる（ステップ 2420）。ライトデータ最古アドレスが、ライトデータ最新アドレスとライトデータ最新アドレスにデータ量 A を足した数値の範囲にある場合、当該ライトデータをライトデータ領域に格納できないため、最古のジャーナルからジャーナルの記憶領域を開放し、ライトデータを格納できるようにする（ステップ 2425）。 40 50

## 【 0 1 1 6 】

( 2 ) ホストアダプタ C は、更新情報をキャッシュメモリ 1 3 0 内に作成する。更新情報のライト命令を受信した時刻は、リモートライト命令内の更新時刻を設定する。グループ番号は、論理ボリューム C の同期ペア番号のペアが属するグループのグループ番号を設定する。更新番号は、リモートライト命令内の更新番号を設定する。ライト命令の論理アドレスは、リモートライト命令内の論理アドレスを設定する。ライトデータのデータサイズは、リモートライト命令内のデータ量 A を設定する。ライトデータを格納したジャーナル論理ボリュームの論理アドレスは、ポインタ情報 7 0 0 のライトデータ最新アドレスであり、ライトデータ最新アドレスにライトデータのサイズを足した数値をポインタ情報 7 0 0 のライトデータ最新アドレスに設定する。更新情報を格納する論理アドレスは、ポインタ情報 7 0 0 の更新情報最新アドレスであり、更新情報のサイズを足した数値をポインタ情報 7 0 0 の更新情報最新アドレスに設定する ( ステップ 2 4 3 0 ) 。

10

## 【 0 1 1 7 】

( 3 ) ホストアダプタ C は、ディスクアダプタ 1 2 0 に、更新情報とライトデータを記憶装置 1 5 0 に書き込むことを命令し、ジャーナル作成成功として処理を終了する ( ステップ 2 4 4 0 ) 。その後、ディスクアダプタ 1 2 0 は、リードライト処理により、記憶装置 1 5 0 に更新情報とライトデータを書き込み、キャッシュメモリ 1 3 0 を開放する。

## 【 0 1 1 8 】

このように、副記憶システム C は、古いジャーナルの記憶領域を開放し、常に新しい複数のジャーナルを保持する。

20

## 【 0 1 1 9 】

リードライト処理 2 2 0 は、ディスクアダプタ 1 2 0 が、ホストアダプタ 1 1 0 もしくはディスクアダプタ 1 2 0 から命令を受け、実施する処理である。実施する処理は、指定されたキャッシュメモリ 1 3 0 のデータを指定された論理アドレスに対応する記憶装置 1 5 0 内の記憶領域に書き込む処理、指定された論理アドレスに対応する記憶装置 1 5 0 内の記憶領域から指定されたキャッシュメモリ 1 3 0 にデータを読み込む処理等である。

## 【 0 1 2 0 】

図 1 4 はジャーナルリード命令を受信した正記憶システム 1 0 0 A のホストアダプタ A の動作 ( ジャーナルリード受信処理 ) を説明する図、図 1 5 はフローチャートである。以下、これらを用いて、正記憶システム 1 0 0 A が、副記憶システム 1 0 0 B からジャーナルリード命令を受信した場合の動作について説明する。

30

## 【 0 1 2 1 】

( 1 ) 正記憶システム 1 0 0 A 内のホストアダプタ A は、副記憶システム 1 0 0 B からアクセス命令を受信する。アクセス命令は、ジャーナルリード命令であることを示す識別子、命令対象のグループ番号、リトライ指示の有無を含んでいる。以下、アクセス命令内のグループ番号をグループ番号 A とする ( ステップ 1 2 2 0 、図 1 4 の 1 4 1 0 ) 。

## 【 0 1 2 2 】

( 2 ) ホストアダプタ A は、グループ番号 A のグループ状態が “ 正常 ” であるかを調べる ( ステップ 1 5 1 0 ) 。ステップ 1 5 1 0 の調べで、グループ状態が “ 正常 ” 以外、例えば、 “ 異常 ” の場合は、副記憶システム 1 0 0 B にグループ状態を通知し、処理を終了する。副記憶システム 1 0 0 B は、受信したグループ状態に応じて処理を行う。例えば、グループ状態が “ 異常 ” の場合は、ジャーナルリード処理を終了する ( ステップ 1 5 1 5 ) 。

40

## 【 0 1 2 3 】

( 3 ) ステップ 1 5 1 0 の調べで、グループ番号 A のグループ状態が “ 正常 ” の場合、ホストアダプタ A は、ジャーナル論理ボリュームの状態を調べる ( ステップ 1 5 2 0 ) 。ステップ 1 5 2 0 の調べで、ジャーナル論理ボリュームのボリューム状態が “ 正常 ” でない場合は、例えば、 “ 異常 ” の場合は、グループ状態を “ 異常 ” に変更し、副記憶システム 1 0 0 B にグループ状態を通知し、処理を終了する。副記憶システム 1 0 0 B は、受信したグループ状態に応じて処理を行う。例えば、グループ状態が “ 異常 ” の場合は、ジャ

50

ーナルリード処理を終了する（ステップ1525）。

【0124】

（4）ステップ1520の調べで、ジャーナル論理ボリュームのボリューム状態が“正常”の場合は、ジャーナルリード命令がリトライ指示かを調べる（ステップ1530）。

【0125】

（5）ステップ1530の調べで、ジャーナルリード命令がリトライ指示の場合、ホストアダプタAは、前回送信したジャーナルを再度、副記憶システム100Bに送信する。ホストアダプタAは、キャッシュメモリ130を確保し、ディスクアダプタに、ポインタ情報700のリトライ開始アドレスから、更新情報のサイズの情報をキャッシュメモリに読み込むことを命令する（図14の1420）。

10

【0126】

ディスクアダプタのリードライト処理は、記憶装置150から更新情報を読み込み、キャッシュメモリ130に保存し、ホストアダプタAに通知する（図14の1430）。

【0127】

ホストアダプタAは、更新情報のリード終了の通知を受け、更新情報から、ライトデータの論理アドレスおよびライトデータのサイズを取得し、キャッシュメモリ130を確保し、ディスクアダプタにライトデータをキャッシュメモリに読み込むことを命令する（ステップ1540、図14の1440）。

【0128】

ディスクアダプタのリードライト処理は、記憶装置150からライトデータを読み込み、キャッシュメモリ130に保存し、ホストアダプタAに通知する（図14の1450）。

20

【0129】

ホストアダプタAは、ライトデータのリード終了の通知を受け、更新情報とライトデータを副記憶システム100Bに送信し、ジャーナルを保持しているキャッシュメモリ130を開放し、処理を終了する（ステップ1545、図14の1460）。

【0130】

（6）ステップ1530の調べで、リトライ指示でない場合、ホストアダプタAは、送信していないジャーナルが存在するかを調べ、存在すれば、ジャーナルを副記憶システム100Bに送信する。ホストアダプタAは、ポインタ情報700のリード開始アドレスと更新情報最新アドレスを比較する（ステップ1550）。

30

【0131】

リード開始アドレスが更新情報最新アドレスと等しい場合は、全てのジャーナルを副記憶システム100Bに送信済みであるため、副記憶システム100Bに“ジャーナル無”を送信し（ステップ1560）、前回のジャーナルリード命令の時に、副記憶システム100Bに送信したジャーナルの記憶領域を開放する（ステップ1590）。

【0132】

ジャーナルの記憶領域の開放処理は、ポインタ情報700の更新情報最古アドレスに、リトライ開始アドレスを設定する。更新情報最古アドレスがライトデータ領域先頭アドレスとなった場合は、更新情報最古アドレスは0とする。ポインタ情報700のライトデータ最古アドレスは、前回のジャーナルリード命令に応じて送信したライトデータのサイズを足した数値に変更する。ライトデータ最古アドレスが、ジャーナル論理ボリュームの容量以上の論理アドレスとなった場合は、ライトデータ領域先頭アドレスを減じ、補正する。

40

【0133】

（7）ステップ1550の調べで、未送信のジャーナルが存在する場合、ホストアダプタAは、キャッシュメモリ130を確保し、ディスクアダプタにポインタ情報700のリード開始アドレスから、更新情報のサイズの情報をキャッシュメモリに読み込むことを命令する（図14の1420）。

【0134】

50

ディスクアダプタ A のリードライト処理は、記憶装置 150 から更新情報を読み込み、キャッシュメモリ 130 に保存し、ホストアダプタに通知する（図 14 の 1430）。

【0135】

ホストアダプタ A は、更新情報のリード終了の通知を受け、更新情報から、ライトデータの論理アドレスおよびライトデータのサイズを取得し、キャッシュメモリ 130 を確保し、ディスクアダプタ A にライトデータをキャッシュメモリに読み込むことを命令する（ステップ 1570、図 14 の 1440）。

【0136】

ディスクアダプタ A のリードライト処理は、記憶装置 150 からライトデータを読み込み、キャッシュメモリ 130 に保存し、ホストアダプタに通知する（図 14 の 1450）

10

【0137】

ホストアダプタ A は、ライトデータのリード終了の通知を受け、更新情報とライトデータを副記憶システム 100 B に送信（ステップ 1580）し、ジャーナルを保持しているキャッシュメモリ 130 を開放する（図 14 の 1460）。そして、ポインタ情報 700 のリトライ開始アドレスにリード開始アドレスを設定し、リード開始アドレスに送信したジャーナルの更新情報サイズを足した数値を設定する。

【0138】

（8）ホストアダプタ A は、前回のジャーナルリード命令の処理時に、副記憶システム 100 B に送信したジャーナルの記憶領域を開放する（ステップ 1590）。

20

【0139】

前述したジャーナルリード受信処理では、正記憶システム 100 A は、ジャーナルを一つずつ副記憶システム 100 B に送信していたが、複数同時に副記憶システム 100 B に送信してもよい。1つのジャーナルリード命令で、送信するジャーナル数は、副記憶システム 100 B がジャーナルリード命令内に指定してもよいし、グループ登録の際等に、ホストコンピュータ 180 は、正記憶システム 100 A もしくは、副記憶システム 100 B に指定してもよい。さらに、正記憶システム 100 A と副記憶システム 100 B の接続バス 200 の転送能力もしくは、負荷等により、動的に1つのジャーナルリード命令で送信するジャーナル数を変更してもよい。また、ジャーナル数でなく、ジャーナルのライトデータのサイズを考慮し、ジャーナルの転送量を指定してもよい。

30

【0140】

前述したジャーナルリード受信処理では、ジャーナルを記憶装置 150 からキャッシュメモリ 130 に読み込んでいたが、キャッシュメモリ 130 に存在する場合は、当該処理は不要である。

【0141】

前述したジャーナルリード受信処理内のジャーナルの記憶領域の開放処理は、次のジャーナルリード命令の処理時としたが、副記憶システム 100 B にジャーナルを送信した直後に開放してもよい。また、副記憶システム 100 B が、ジャーナルリード命令内に開放してよい更新番号を設定し、正記憶システム 100 A は、その指示に従って、ジャーナルの記憶領域を開放してもよい。

40

【0142】

図 16 はジャーナルリード処理 240 を説明する図、図 17 はフローチャート、図 18 はジャーナル格納処理のフローチャートである。以下、これらを用いて、副記憶システム 100 B のホストアダプタ B が、正記憶システム 100 A からジャーナルを読み出し、ジャーナル論理ボリュームに格納する動作について説明する。

【0143】

（1）副記憶システム 100 B 内のホストアダプタ B は、グループ状態が“正常”かつ複製種類が非同期であれば、ジャーナルを格納するキャッシュメモリ 130 を確保し、ジャーナルリード命令であることを示す識別子、命令対象の正記憶システム 100 A のグループ番号、リトライ指示の有無を含むアクセス命令を正記憶システム 100 A に送信する

50

。以下、アクセス命令内のグループ番号をグループ番号Aとする（ステップ1700、図16の1610）。

【0144】

（2）ホストアダプタBは、正記憶システム100Aの応答およびジャーナルを受信する（図16の1620）。ホストアダプタBは応答を調べ、正記憶システム100Aからの応答が、“ジャーナル無”の場合は、正記憶システム100Aには、指定したグループのジャーナルが存在しないため、一定時間後、正記憶システム100Aにジャーナルリード命令を送信する（ステップ1720、1725）。

【0145】

（4）正記憶システム100Aの応答が、“グループ状態は異常”もしくは“グループ状態は未使用”の場合は、副記憶システム100Bのグループ状態を受信した状態に変更し、ジャーナルリード処理を終了する（ステップ1730、1735）。

10

【0146】

（5）正記憶システム100Aの応答が、上記以外、つまり、正常終了の場合は、ジャーナル論理ボリュームのボリューム状態を調べる（ステップ1740）。ジャーナル論理ボリュームのボリューム状態が“異常”の場合は、ジャーナル論理ボリュームにジャーナルの格納が不可能なため、グループ状態を“異常”に変更し、処理を終了する（ステップ1745）。この場合、ジャーナル論理ボリュームを正常な論理ボリュームに変更する等を行い、グループの状態を正常に戻す。

【0147】

20

（6）ステップ1740の調べで、ジャーナル論理ボリュームのボリューム状態が“正常”の場合は、後述するジャーナル格納処理1800を行う。ジャーナル格納処理1800が正常に終了した場合は、次のジャーナルリード命令を送信する。もしくは、一定時間経過後、次のジャーナルリード命令を送信する（ステップ1700）。次のジャーナル命令を送信するタイミングは、一定の時間間隔で定期的に送信するものでもよく、受信したジャーナルの個数もしくは、接続線200の通信量、副記憶システム100Bが保持しているジャーナルの記憶容量、副記憶システム100Bの負荷等によって決めてもよい。さらに、正記憶システム100Aが保持しているジャーナルの記憶容量、もしくは正記憶システム100Aのポインタ情報を副記憶システム100Bから読み出し、その数値に基づいて決めてもよい。上記情報の転送は、専用のコマンドで行ってもよいし、ジャーナルリード命令の応答に含んでもよい。その後の処理は、ステップ1700以降と同じである。

30

【0148】

（7）ステップ1800のジャーナル格納処理が正常に終了しない場合は、ジャーナル論理ボリュームの未使用領域が足りないため、受信したジャーナルを破棄し、一定時間後にリトライ指示のジャーナルリード命令を送信する（ステップ1755）。もしくは、ジャーナルをキャッシュメモリに保持しておき、一定時間後に、再度ジャーナル格納処理を行う。これは、後述するリストア処理250が行われることにより、一定時間後には、ジャーナル論理ボリュームに未使用領域が増える可能性があるためである。この方式の場合は、ジャーナルリード命令にリトライ指示の有無は不要である。

【0149】

40

次に、図18に示すジャーナル格納処理1800について説明する。

（1）ホストアダプタBは、ジャーナルがジャーナル論理ボリュームに格納可能であることを調べる。ポインタ情報700を用い、更新情報領域に未使用領域の有無を調べる（ステップ1810）。ポインタ情報700の更新情報最新アドレスと更新情報最古アドレスが等しい場合は、更新情報領域に未使用領域が存在しないため、ジャーナル作成失敗として処理を終了する（ステップ1820）。

【0150】

（2）ステップ1810の調べで、更新情報領域に未使用領域が存在する場合は、ポインタ情報700を用い、ライトデータ領域にライトデータが格納できるかを調べる（ステップ1830）。ライトデータ最古アドレスが、ライトデータ最新アドレスとライトデー

50

タ最新アドレスにデータ量 A を足した数値の範囲にある場合、ライトデータ領域にライトデータを格納できないため、ジャーナル作成失敗として処理を終了する（ステップ 1820）。

【0151】

（3）ジャーナルが格納可能である場合、ホストアダプタ B は、受信した更新情報のグループ番号とライトデータを格納したジャーナル論理ボリュームの論理アドレスを変更する。グループ番号は、副記憶システム 100B のグループ番号に変更し、ジャーナル論理ボリュームの論理アドレスはポインタ情報 700 のライトデータ最新アドレスに変更する。さらに、ホストアダプタ B は、ポインタ情報 700 の更新情報最新アドレスを更新情報最新アドレスに更新情報のサイズを足した数値に、ライトデータ最新アドレスを、ライトデータ最新アドレスにライトデータのサイズを足した数値に変更する。さらに、ホストアダプタ B は、グループ情報の更新番号を受信した更新情報の更新番号に変更する（ステップ 1840）。

10

【0152】

（4）ホストアダプタ B は、ディスクアダプタ 120 に、更新情報とライトデータを記憶装置 150 に書き込むことを命令し、ジャーナル作成成功として処理を終了する（ステップ 1850、図 16 の 1630）。その後、ディスクアダプタ 120 は、リードライト処理により、記憶装置 150 に更新情報とライトデータを書き込み、キャッシュメモリ 130 を開放する（図 16 の 1640）。

【0153】

20

前述したジャーナル格納処理では、ジャーナルを記憶装置 150 に保存するとしていたが、ジャーナル用に予め一定量のキャッシュメモリ 130 を用意しておき、当該キャッシュメモリを全て使用してから、記憶装置 150 にジャーナルを保存してもよい。ジャーナル用のキャッシュメモリ量は、例えば、保守端末から指定する。

【0154】

図 19 はリストア処理 250 を説明する図、図 20 はフローチャートである。以下、これらを用いて、副記憶システム 100B のホストアダプタ B が、ジャーナルを利用し、データの更新を行う動作について説明する。リストア処理 250 は副記憶システム 100B のディスクアダプタ 120 が処理を行ってもよい。

【0155】

30

（1）ホストアダプタ B は、グループ番号 B のグループ状態が“正常”もしくは“停止”であるかを調べる（ステップ 2010）。ステップ 2010 の調べで、グループ状態が“正常”および“停止”以外、例えば、“異常”の場合は、リストア処理を終了する（ステップ 2015）。

【0156】

（2）ステップ 2010 の調べで、グループ状態が“正常”もしくは“停止”の場合は、ジャーナル論理ボリュームのボリューム状態を調べる（ステップ 2020）。ステップ 2020 の調べで、ジャーナル論理ボリュームのボリューム状態が、“異常”の場合は、アクセス不可能なため、グループ状態を“異常”に変更し、処理を終了する（ステップ 2025）。

40

【0157】

（3）ステップ 2020 の調べで、ジャーナル論理ボリュームのボリューム状態が、“正常”の場合は、リストア対象のジャーナルが存在するかを調べる。ホストアダプタ B は、ポインタ情報 700 の更新情報最古アドレスと更新情報最新アドレスを取得する。更新情報最古アドレスと更新情報最新アドレスが等しい場合、ジャーナルは存在しないため、リストア処理は一旦終了し、一定時間後、リストア処理を再開する（ステップ 2030）。

【0158】

（4）ステップ 2030 の調べで、リストア対象のジャーナルが存在する場合、最古（最小）の更新番号を持つジャーナルに対して次の処理を行う。最古（最小）の更新番号を

50

持つジャーナルの更新情報は、ポインタ情報 700 の更新情報最古アドレスから保存されている。ホストアダプタ B は、キャッシュメモリ 130 を確保し、ディスクアダプタに更新情報最古アドレスから、更新情報のサイズの情報をキャッシュメモリ 130 に読み込むことを命令する（図 19 の 1910）。

【0159】

ディスクアダプタのリードライト処理は、記憶装置 150 から更新情報を読み込み、キャッシュメモリ 130 に保存し、ホストアダプタ B に通知する（図 19 の 1920）。

【0160】

ホストアダプタ B は、更新情報のリード終了の通知を受け、更新情報から、ライトデータの論理アドレスおよびライトデータのサイズを取得し、キャッシュメモリ 130 を確保し、ディスクアダプタにライトデータをキャッシュメモリに読み込むことを命令する（図 19 の 1930）。

10

【0161】

ディスクアダプタのリードライト処理は、記憶装置 150 からライトデータを読み込み、キャッシュメモリ 130 に保存し、ホストアダプタに通知する（ステップ 2040、図 19 の 1940）。

【0162】

（5）ホストアダプタ B は、更新情報から更新する副論理ボリュームの論理アドレスを求め、ディスクアダプタに副論理ボリュームにライトデータを書き込むことを命令する（ステップ 2050、図 19 の 1950）。ディスクアダプタのリードライト処理は、副論理ボリュームの論理アドレスに対応する記憶装置 150 にデータを書き込み、キャッシュメモリ 130 を開放し、ホストアダプタに通知する（図 19 の 1960）。

20

【0163】

（6）ホストアダプタ B は、ディスクアダプタのライト処理完了の通知を受け、ジャーナルの記憶領域を開放する。ジャーナルの記憶領域の開放処理は、ポインタ情報 700 の更新情報最古アドレスを更新情報のサイズを足した数値に変更する。更新情報最古アドレスが、ライトデータ領域先頭アドレスとなった場合は、更新情報最古アドレスは 0 とする。ポインタ情報 700 のライトデータ最古アドレスは、ライトデータのサイズを足した数値に変更する。ライトデータ最古アドレスが、ジャーナル論理ボリュームの容量以上の論理アドレスとなった場合は、ライトデータ領域先頭アドレスを減じ、補正する。その後、ホストアダプタ B は、次のリストア処理を開始する（ステップ 2060）。

30

【0164】

前述したリストア処理 250 では、記憶装置 150 からキャッシュメモリ 130 にジャーナルを読み込んでいたが、キャッシュメモリ 130 に存在する場合は、当該処理は不要である。

【0165】

前述したジャーナルリード受信処理とジャーナルリード処理 240 では、正記憶システム 100 A が送信するジャーナルをポインタ情報 700 により決めていたが、副記憶システム 100 B が送信するジャーナルを決めてもよい。例えば、ジャーナルリード命令に更新番号を追加する。この場合、ジャーナルリード受信処理にて、副記憶システム 100 B が指定した更新番号の更新情報の論理アドレスを求めるために、正記憶システム 100 A の共有メモリ 140 内に、更新番号から更新情報を格納した論理アドレスを求めるテーブルもしくは検索方法を設ける。

40

【0166】

前述したジャーナルリード受信処理とジャーナルリード処理 240 では、ジャーナルリード命令を用いていたが、通常のリード命令を用いてもよい。例えば、正記憶システム 100 A のグループ情報 600 とポインタ情報 700 を予め副記憶システム 100 B に転送しておき、副記憶システム 100 B は、正記憶システム 100 A のジャーナル論理ボリュームのデータ（つまり、ジャーナル）をリードする。

【0167】

50

前述したジャーナルリード受信処理では、更新番号の順に、正記憶システム100Aから副記憶システム100Bにジャーナルを送信すると説明したが、更新番号の順に送信せずともよい。また、正記憶システム100Aから副記憶システム100Bに複数のジャーナルリード命令を送信してもよい。この場合、リストア処理にて更新番号順にジャーナルを処理するために、副記憶システム100Bに、更新番号から更新情報を格納した論理アドレスを求めるテーブルもしくは検索方法を設ける。

#### 【0168】

前述した本発明の計算機システムでは、記憶システムAは、データの更新に関する情報をジャーナルとして格納する。記憶システムBは、記憶システムAが保持するデータの複製を保持しており、自律的に記憶システムAからジャーナルを取得し、前記ジャーナルを用いて、第一の記憶システムでのデータ更新順に、第一の記憶システムのデータと対応するデータを更新する。これにより、記憶システムBは、データの整合性を維持しつつ、記憶システムAのデータの複製を保持できる。さらに、ジャーナルを管理する管理情報は、複製対象のデータ容量に依存しない。

#### 【0169】

正記憶システム100Aが故障した場合に、ホストコンピュータ180にて行っている情報処理を、ホストコンピュータ180Cおよび記憶システム100Cを用いて再開し、記憶システム100Bでのデータ複製を再開する手順を図25に、論理的な構成を示すブロック図を図42に示す。ホストコンピュータ180とホストコンピュータ180Cは同じコンピュータであってもよい。

#### 【0170】

以下の説明において、正記憶システム100Aが故障する前の正記憶システム100Aのボリューム情報は図4、ペア情報は図5、グループ情報は図6、ポインタ情報は図7、ポインタ情報を説明する図は図8とする。正記憶システム100Aが故障する前の副記憶システム100B（非同期複製）のボリューム情報は図26、ペア情報は図27、グループ情報は図28、ポインタ情報は図29、ポインタ情報を説明する図は図30とする。副記憶システム100Bは、非同期のデータ複製を行っているため、正記憶システム100Aが保持するジャーナル（更新番号3～5）の全てを保持していない場合がある。本例では、副記憶システム100Bは、更新番号5のジャーナルを保持していない。正記憶システム100Aが故障する前の副記憶システム100C（同期複製）のボリューム情報は図31、ペア情報は図32、グループ情報は図33、ポインタ情報は図34、ポインタ情報を説明する図は図35とする。副記憶システム100Cは、同期のデータ複製を行っているため、正記憶システム100Aが保持する全てのジャーナル（更新番号3～5）を保持する。

#### 【0171】

（1）正記憶システム100Aに故障が発生し、正論理ボリューム（DATA1、DATA2等）が使用不可能となる（ステップ2500）。

#### 【0172】

（2）ホストコンピュータ180は、記憶システム100Cに、デルタリシンクを命令する。デルタリシンク命令は、グループ単位で、非同期のデータ複製元（正論理ボリューム）を変更する命令であり、複製元情報（同期のデータ複製の副論理ボリューム（data1、data2）を保持する記憶システム番号Cとグループ番号D）、複製先情報（非同期のデータ複製の副論理ボリューム（COPY1、COPY2）を保持する記憶システム番号Bとグループ番号B）を含む。グループ番号Dとグループ番号Bは、双方ともデルタ予約オプションが“1”である（ステップ2510）。

#### 【0173】

（3）記憶システム100Cは、デルタリシンク命令を受信すると、記憶システム100Bおよび記憶システム100Cのボリューム情報、ペア情報、グループ情報を参照し、記憶システム100C内の既存の同期リモートコピーのグループであるグループ1に属する論理ボリュームのペア状態を“正常”から“異常”に変更する。また、記憶システム1

00C内のグループ2に属する論理ボリュームのペア状態を“デルタ待機”から“正常”に変更する。さらに、記憶システム100Cは、グループ1に属していたジャーナル論理ボリュームをグループ2にて継続して使用するようグループ情報を変更する。具体的には、記憶システム100Cは、グループ2の更新番号をグループ1の更新番号に、グループ2のジャーナル論理ボリューム番号をグループ1のジャーナル論理ボリューム番号に変更し、グループ2のポインタ情報の全ての項目を、グループ1のポインタ情報と同じにする。デルタリシンク命令により、記憶システム100Cは、図32に示す記憶システム100Cのペア情報を図39に示すペア情報に、図33に示す記憶システム100Cのグループ情報を図40に示すグループ情報に、図31に示す記憶システム100Cのボリューム情報を図38に示すボリューム情報に変更する。

10

#### 【0174】

記憶システム100Cは、記憶システム100Bに対して、記憶システム100C内のグループ1に属する論理ボリュームのペア状態を“正常”から“異常”に変更する。また、同様に記憶システム100C内のグループ2に属する論理ボリュームのペア状態を“デルタ待機”から“正常”に変更する。

#### 【0175】

記憶システム100Bは、記憶システム100Bおよび記憶システム100Cのボリューム情報、ペア情報、グループ情報を参照し、記憶システム100Bのペア情報、グループ情報を変更する。記憶システム100Bは、図27に示す記憶システム100Bのペア情報を図36に示すペア情報に、図28に示すグループ情報を図37に示すグループ情報に、グループ1のグループ情報の状態を“停止”に変更することで、記憶システム100Aへのジャーナルリード処理を停止する。(ステップ2530、2540)。

20

#### 【0176】

(4) 記憶システム100Cは、デルタリシンク命令の応答をホストコンピュータ180もしくは保守端末に行う。ホストコンピュータ180は、記憶システム100Cからデルタリシンク命令の応答を受領し、デルタリシンクの終了を認識し、記憶システム100Cの使用を開始する(ステップ2550、2560)。

#### 【0177】

(5) 記憶システム100Bは、記憶システム100Cにジャーナルリード位置指定命令を送信する(ステップ2570)。ジャーナルリード位置指定命令は、記憶システム100Cのグループ2のポインタ情報を変更し、記憶システム100Bからのジャーナルリード命令により送信されるジャーナルを指定する命令であり、相手グループ番号、更新番号を含む。相手グループ番号は、グループ番号2の相手グループ番号を指定する。更新番号は、グループ番号2のグループ情報の更新番号に1を足した数値を指定する。図37に示す例では、グループ番号2、更新番号5を指定する。

30

#### 【0178】

(6) 記憶システム100Cは、ジャーナルリード位置指定命令を受信した場合、ポインタ情報700を参照し、更新番号Bのジャーナルを保持しているかを調べる。記憶システム100Cは、ポインタ情報の更新情報最古アドレスの更新情報を記憶装置150から読み出し、最古(最小)の更新番号Cを得る。

40

#### 【0179】

更新番号Cがジャーナルリード位置指定命令の更新番号Bと等しい、もしくは小さい場合は、記憶システム100Cが更新番号Bのジャーナルを保持しているので、記憶システム100Bは、非同期のデータの複製を継続することができる。この場合、記憶システム100Cは、更新番号Bより以前のジャーナルの記憶領域を開放し、リード開始アドレスおよびリトライ開始アドレスを更新番号Bの更新情報を格納しているアドレスに変更し、“再開可能”を記憶システム100Bに返す。これにより、図34に示すポインタ情報は、図41に示すポインタ情報に変更される(ステップ2580)。

#### 【0180】

他方、更新番号Cがジャーナルリード位置指定命令の更新番号Bより大きい場合は、記

50

憶システム 100C が記憶システム 100B に必要なジャーナルを保持していないため、記憶システム 100B において非同期のデータの複製を継続することができない。この場合は、図 9、図 10 を用いて説明した手順により、正記憶システム 100C から副記憶システム 100B に対して、データ複製を開始する必要がある。

【0181】

(7) 記憶システム 100B は、“再開可能”の応答を受信した場合、グループ B のグループ情報の状態を“正常”に変更することで、記憶システム 100C に対し、ジャーナルリード処理を再開する(ステップ 2590)。

【0182】

記憶システム 100B は、ジャーナルリード位置指定命令を行わなくてもよい。この場合、記憶システム 100B は、ジャーナルリード処理を開始し、記憶システム 100C から最古のジャーナルを受信する。受信したジャーナルの更新番号 C が、グループ番号 B のグループ情報の更新番号に 1 を足した数値(更新番号 B)より大きい場合は、記憶システム 100C が記憶システム 100B に必要なジャーナルを保持していないため、データ複製処理を中断する。受信したジャーナルの更新番号 C が、更新番号 B より小さい場合は、すでに記憶システム 100B が当該ジャーナルを保持しているため、当該ジャーナルを破棄し、ジャーナルリード処理を継続する。受信したジャーナルの更新番号 C が、更新番号 B と等しい場合は、受信したジャーナルをジャーナル論理ボリュームに格納し、ジャーナルリード処理を継続する。

【0183】

図 42 を用いて、ホストコンピュータ 180C が記憶システム 100C の使用を開始した後の、正記憶システム 100C の正論理ボリューム(data1)へのデータ更新を副記憶システム 100B の副論理ボリューム(COPY1)に反映する動作について概説する。

【0184】

(1) 正記憶システム 100C は、ホストコンピュータ 180C から正論理ボリューム(data1)内のデータに対するライト命令を受信すると、前述した命令受信処理 210 およびリードライト処理 220 によって、正論理ボリューム(data1)内のデータ更新と、ジャーナル論理ボリューム(jnl1)へのジャーナルの保存を行い、ホストコンピュータにライト命令の終了を報告する(図 42 の 4200)。

【0185】

(2) 副記憶システム 100B は、前述したジャーナルリード処理 240 によって、正記憶システム 100C からジャーナルをリードし、リードライト処理 220 によって、ジャーナル論理ボリューム(jnl2)にジャーナルを保存する(図 42 の 4210)。

【0186】

(3) 正記憶システム 100C は、副記憶システム 100B からジャーナルリード命令を受信すると、命令受信処理 210 およびリードライト処理 220 によって、ジャーナル論理ボリューム(jnl1)からジャーナルを読み出し、副記憶システム 100B に送信する(図 42 の 4210)。

【0187】

(4) 副記憶システム 100B は、リストア処理 250 およびリードライト処理 220 によって、ポインタ情報 700 を用いて、更新番号の昇順に、ジャーナル論理ボリューム(jnl2)からジャーナルを読み出し、副論理ボリューム(COPY1)のデータを更新する(図 42 の 4220)。従って、正記憶システム 100C 内の正論理ボリューム(data1)と副記憶システム 100B 内の副論理ボリューム(COPY1)のデータは、正論理ボリュームの更新の暫く後には、完全に一致する。

【0188】

次に、デルタリシンク用のグループ及びペアを生成した後に、ジャーナル論理ボリュームの障害や、記憶システム C と記憶システム B の間の接続パス 200 の切断などの障害により、デルタリシンク用のペア状態が“デルタ異常”に遷移した後の回復手段であるデル

10

20

30

40

50

タ回復処理 2 6 0 0 について図 4 3 を用いて説明を行う。

【 0 1 8 9 】

ユーザはペア状態を“デルタ異常”として検知した後、まず障害を取り除く。その後、ホストコンピュータ 1 8 0 は、デルタ回復処理指示を行う。デルタ回復処理指示は、指示対象グループ番号 D、相手記憶システム番号 B、相手グループ番号 B からなる。

【 0 1 9 0 】

まず、副記憶システム 1 0 0 C は、指示対象グループ番号 D の論理ボリュームに対応するグループ番号 C に対し、ジャーナル複製処理を再開する（ステップ 2 6 1 0 ）。

【 0 1 9 1 】

次に、副記憶システム 1 0 0 C は、“デルタ異常”になっているグループ番号 D の論理ボリュームのペア状態を、“デルタ準備”に変更する（ステップ 2 6 2 0 ）。

10

【 0 1 9 2 】

次に、副記憶システム 1 0 0 C は、前述のデルタ予約処理と同様の、デルタリシンクが可能か否かのチェックを行う（ステップ 2 6 3 0 ）。すなわち、まずデルタリシンク用に登録したペアを複製先としている 2 つのグループの複製対象が一致し、かつ、ペア数も一致しているかを、副記憶システム 1 0 0 C は、次のような手順で比較し、判定する。

【 0 1 9 3 】

副記憶システム 1 0 0 C は、ペア情報 5 0 0 の中から、指定されたグループ番号 D であるペアの正記憶システム番号 C と正論理ボリューム番号 C と同一である副記憶システム番号と副論理ボリューム番号を持つペア情報を抽出し、そこから正記憶システム番号 A 2 と正論理ボリューム番号 A 1 を取得する（ステップ 4 6 3 0 ）。次に、副記憶システム 1 0 0 C は、正記憶システム番号 C と正論理ボリューム番号 C に対応する副記憶システム番号 B である副記憶システム 1 0 0 B から、正記憶システム番号 C と正論理ボリューム番号 C に対応する副記憶システム番号 B と副論理ボリューム B を複製先として有し、デルタ予約オプションが“ 1 ”でないグループ番号もつペアの正記憶システム番号 A 2 と、正論理ボリューム番号 A 2 を取得する（ステップ 4 6 4 0 ）。そして副記憶システム 1 0 0 C は、取得した正記憶システム番号 A 1 及び正論理ボリューム番号 A 1 と、正記憶システム番号 A 2 と正論理ボリューム番号 A 2 が一致するかを判定する（ステップ 4 6 5 0 ）。この判定を、ステップ 4 6 6 0、ステップ 4 6 2 0 を経て指定されたグループ番号 C であるペア全てに対して行う。

20

30

【 0 1 9 4 】

この判定で不一致がでた場合は、副記憶システム 1 0 0 C は、エラーで完了する。

【 0 1 9 5 】

次のチェックとして、副記憶システム 1 0 0 C は、デルタリシンクを行う場合に必要なジャーナルデータの転送が行われているかを比較し判定する。

【 0 1 9 6 】

まず、副記憶システム 1 0 0 C は、グループ情報 6 0 0 の中で、指定されたグループ番号 D でデルタ予約オプションが“ 1 ”であるグループ番号に記載されている相手記憶システム番号 B である副記憶システム 1 0 0 B から、当該グループ番号に記載されている相手グループ番号 B であるグループの更新番号 B に 1 を足した数値 B 1 を取得する（ステップ 4 6 7 2 ）。次に、記憶システム 1 0 0 C は、ポイント情報 7 0 0 を参照し、更新番号 B のジャーナルを保持しているかを調べる。記憶システム 1 0 0 C は、ポイント情報の更新情報最古アドレスの更新情報を記憶装置 1 5 0 から読み出し、最古（最小）の更新番号 C を得る（ステップ 4 6 7 4 ）。更新番号 C が取得した更新番号 B 1 と比較し（ステップ 4 6 7 6 ）、更新番号 C が更新番号 B 1 と等しい、もしくは更新番号 B 1 より小さい場合は、記憶システム 1 0 0 C が更新番号 B のジャーナルを保持しているので、デルタリシンクが可能な状態と判定する（ステップ 4 6 7 8 ）。

40

【 0 1 9 7 】

デルタリシンクが可能な状態と判定した場合は、当該グループ番号のペア情報 5 0 0 のペア状態を“デルタ待機”に変更し、処理を終了する。

50

## 【 0 1 9 8 】

前述の処理においては、ホストコンピュータ 1 8 0 が記憶システムに対して行う指示は、上述のストレージ管理プログラムが提供するユーザインタフェースを介してユーザの操作による入力を契機で行っていることを意味する。この構成を図 4 4 を用いて説明する。図 4 4 では計算機システム 1 0 0 A と、それに接続されているホストコンピュータ 1 8 0 を用いて説明するが、記憶システム 1 0 0 B、及び、記憶システム 1 0 0 C についても同様である。

## 【 0 1 9 9 】

ホストコンピュータ 1 8 0 は、CPU 1 8 0 - 1、主記憶装置 1 8 0 - 2、インターフェース ( I / F ) 1 8 0 - 3 及び管理 I / F 1 8 0 - 4 によって構成される。また、ホストコンピュータ 1 8 0 にはホストコンピュータサービスプロセッサ ( S V P ) 1 8 0 - 5 が接続される。

10

## 【 0 2 0 0 】

CPU 1 8 0 - 1 は、ホストコンピュータ 1 8 0 における処理装置である。すなわち、主記憶装置 1 8 0 - 2 に格納されている OS やプログラム等を読み込んで実行し、その OS やプログラムに規定された処理を実行する。

## 【 0 2 0 1 】

主記憶装置 1 8 0 - 2 は、DRAM 等のメモリ装置によって構成される。この主記憶装置 1 8 0 - 2 には、OS 1 8 0 - 6、ストレージ管理プログラム 1 8 0 - 7 が格納される。

20

## 【 0 2 0 2 】

OS 1 8 0 - 6 は、ホストコンピュータ 1 8 0 のオペレーティングシステムである。OS 1 8 0 - 6 が CPU 1 8 0 - 1 によって実行されることで、ホストコンピュータ 1 8 0 の基本的な動作をする。

## 【 0 2 0 3 】

ストレージ管理プログラム 1 8 0 - 7 は、記憶システム 1 0 0 の管理を行うため、計算機システム 1 0 0 に制御コマンドを発行し、また、ユーザインタフェースを提供するプログラムで、CPU 1 8 0 - 1 に読み込まれて実行される。

## 【 0 2 0 4 】

I / F 1 8 0 - 3 は、記憶システム 1 0 0 とデータを送受信するインターフェースである。

30

## 【 0 2 0 5 】

ホストコンピュータ S V P 1 8 0 - 5 は、管理者ユーザ等の指示によって、ホストコンピュータ 1 8 0 にメンテナンス等の指示を入力する入出力用制御端末であり、ホストコンピュータ 1 8 0 はホストコンピュータ S V P 1 8 0 - 5 とデータを送受信するインターフェースである管理 I / F 1 8 0 - 4 を持つ。

## 【 0 2 0 6 】

ストレージ管理プログラム 1 8 0 - 7 は、実施形態で説明したデルタリシンク用ペア作成及び状態情報表示を含むストレージ管理を行う。

## 【 0 2 0 7 】

命令受信処理 2 1 0 は、実施形態で説明したデルタリシンク用のペアの予約処理、ペア状態管理を含む命令受信処理を行う。

40

## 【 0 2 0 8 】

ペア情報 5 0 0 は、実施形態で説明したデルタリシンク用の予約ペアの状態情報を含むペア情報をである。ペア情報 5 0 0 の詳細は、例えば図 5、図 2 7、図 3 2 である。

## 【 0 2 0 9 】

グループ情報 6 0 0 は、実施形態で説明したデルタリシンク用の予約ペアのためのオプション情報を含むグループ情報をである。グループ情報 6 0 0 の詳細は、例えば図 6、図 2 8、図 3 3 である。

## 【 0 2 1 0 】

50

また、ホストコンピュータ180は、ユーザのデータ入力を容易にするためにグラフィカルユーザインタフェース（GUI）を有することとしてもよい。このGUIは、例えばホストコンピュータ180に接続される記憶システム100A、100B、100Cをその接続トポロジーの通りにホストコンピュータSVP180-5に表示する。そして、GUIは、ユーザからの指示を受け付けて、指示された記憶システム100A、100B、100Cにおいて、リモートコピーペアの管理コマンド情報を作成してストレージ管理プログラム180-7に指示することとしてもよい。また、GUIを用いて、記憶システム100A、100B、100C間のデルタリシンク用ペアを含むペアやグループの情報をホストコンピュータSVP180-5に視覚的に表示してもよい。これにより、ユーザは、個々のペアやグループをGUIのインタフェースにより選択することができ、容易にデルタリシンクの運用指示をホストコンピュータ180に指示することができる。また、記憶システム100Bと記憶システム100Cの間のデルタリシンク用のペア情報を、記憶システム100A及記憶システムB間のペア情報、及び、記憶システム100A及び記憶システム100Cの間のペア情報を参照する形態で指示してもよい。これにより、デルタ予約処理時のエラー発生を低減することができる。この時のGUIの例を図45に示す。

10

#### 【0211】

前述した実施形態の計算機システムでは、記憶システムCが記憶システムAからの更新番号および更新時刻を用いて、ジャーナルを作成する。データ複製対象の記憶システムAが故障し、記憶システムCを用いて情報処理を継続する場合に、記憶システムBは、記憶システムAから記憶システムCにジャーナルの取得先を変更する。これにより、記憶システムBは、データの整合性を維持しつつ、記憶システムAのデータの複製を継続できる。さらに、ジャーナルを管理する管理情報は、複製対象のデータ容量に依存しない。

20

#### 【0212】

前述した本発明の計算機システムでは、デルタリシンク用のペアの状態を、記憶システムの状態に基づき、デルタリシンク可能な“デルタ待機”、障害などのユーザ操作を必要とする“デルタ異常”、デルタリシンク用ペアの登録処理中や、“デルタ異常”からの回復中で、デルタリシンク不可な“デルタ準備”の3つの状態を用いてユーザに表示を行う。また、デルタリシンク用のペアを、デルタリシンクを行う前に予め専用のペアとして登録するデルタ予約処理手順をユーザに提供する。更に、“デルタ準備”及び“デルタ異常”から“デルタ待機”へのペア状態遷移を指示するデルタ回復処理手順をユーザに提供する。これによりユーザは、デルタリシンク用ペアの管理を、デルタリシンク指示前に行うことが可能になる。

30

#### 【0213】

以上に述べた通り、本実施形態での計算機システムでは、記憶システムAもしくは記憶システムAを用いて処理を行っている計算機の故障や保守のための計画停止の際に、記憶システムCを用いて業務を継続するために、記憶システムCと記憶システムBの間でデータの整合性を維持しつつ差分に相当するデータをコピーすることで同一性を確保するデルタリシンクを行う場合に、記憶システムCと記憶システムBにおけるデータ転送設定を予めホストコンピュータから監視することができ、デルタリシンクできない状態を解消することができる。その結果、デルタリシンクできる状態の期間を延ばすことができ、記憶システムAに障害が発生した場合も、直ちに、記憶システムCがデルタリシンクを行い、業務を引き継ぐことができる。以上、本発明者によってなされた発明を実施例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。

40

#### 【図面の簡単な説明】

#### 【0214】

【図1】本発明の一実施形態の論理的な構成を示すブロック図である。

【図2】本発明の一実施形態の記憶システムのブロック図である。

【図3】本発明の一実施形態の更新情報とライトデータの関係を示す図である。

50

【図 4】本発明の一実施形態のボリューム情報の例を説明する図である。

【図 5】本発明の一実施形態のペア情報の例を説明する図である。

【図 6】本発明の一実施形態のグループ情報の例を説明する図である。

【図 7】本発明の一実施形態のポインタ情報の例を説明する図である。

【図 8】本発明の一実施形態のジャーナル論理ボリュームの構造を説明する図である。

【図 9】本発明の一実施形態のデータの複製を開始する手順を説明するフローチャートである。

【図 10】本発明の一実施形態の初期コピー処理を説明するフローチャートである。

【図 11】本発明の一実施形態の命令受信処理を説明する図である。

【図 12】本発明の一実施形態の命令受信処理のフローチャートである。

10

【図 13】本発明の一実施形態のジャーナル作成処理のフローチャートである。

【図 14】本発明の一実施形態のジャーナルリード受信処理を説明する図である。

【図 15】本発明の一実施形態のジャーナルリード受信処理のフローチャートである。

【図 16】本発明の一実施形態のジャーナルリード処理を説明する図である。

【図 17】本発明の一実施形態のジャーナルリード処理のフローチャートである。

【図 18】本発明の一実施形態のジャーナル格納処理のフローチャートである。

【図 19】本発明の一実施形態のリストア処理を説明する図である。

【図 20】本発明の一実施形態のリストア処理のフローチャートである。

【図 21】本発明の一実施形態の更新情報の例を説明する図である。

【図 22】本発明の一実施形態のジャーナル作成処理時の更新情報の例を説明する図である。

20

【図 23】本発明の一実施形態のリモートライト命令受信処理のフローチャートである。

【図 24】本発明の一実施形態のジャーナル複製処理のフローチャートである。

【図 25】本発明の一実施形態の正記憶システム 100A が故障した場合に記憶システム間のデータ複製を再開する手順を説明するフローチャートである。

【図 26】本発明の一実施形態のボリューム情報の例を説明する図である。

【図 27】本発明の一実施形態のペア情報の例を説明する図である。

【図 28】本発明の一実施形態のグループ情報の例を説明する図である。

【図 29】本発明の一実施形態のポインタ情報の例を説明する図である。

【図 30】本発明の一実施形態のジャーナル論理ボリュームの構造を説明する図である。

30

【図 31】本発明の一実施形態のボリューム情報の例を説明する図である。

【図 32】本発明の一実施形態のペア情報の例を説明する図である。

【図 33】本発明の一実施形態のグループ情報の例を説明する図である。

【図 34】本発明の一実施形態のポインタ情報の例を説明する図である。

【図 35】本発明の一実施形態のジャーナル論理ボリュームの構造を説明する図である。

【図 36】本発明の一実施形態のペア情報の例を説明する図である。

【図 37】本発明の一実施形態のグループ情報の例を説明する図である。

【図 38】本発明の一実施形態のボリューム情報の例を説明する図である。

【図 39】本発明の一実施形態のペア情報の例を説明する図である。

【図 40】本発明の一実施形態のグループ情報の例を説明する図である。

40

【図 41】本発明の一実施形態のポインタ情報の例を説明する図である。

【図 42】本発明の一実施形態の正記憶システム 100A が故障した場合の動作を説明するブロック図である。

【図 43】本発明の一実施形態のデルタ回復処理のフローチャートである。

【図 44】本発明の一実施形態のホストコンピュータ 180 とデルタリシンクに関連する部位の論理的な構成を示すブロック図である。

【図 45】本発明の一実施形態の G U I の例を説明する図である。

【図 46】本発明の一実施形態のデルタリシンクが可能か否かを判定する処理を説明する図である。

【符号の説明】

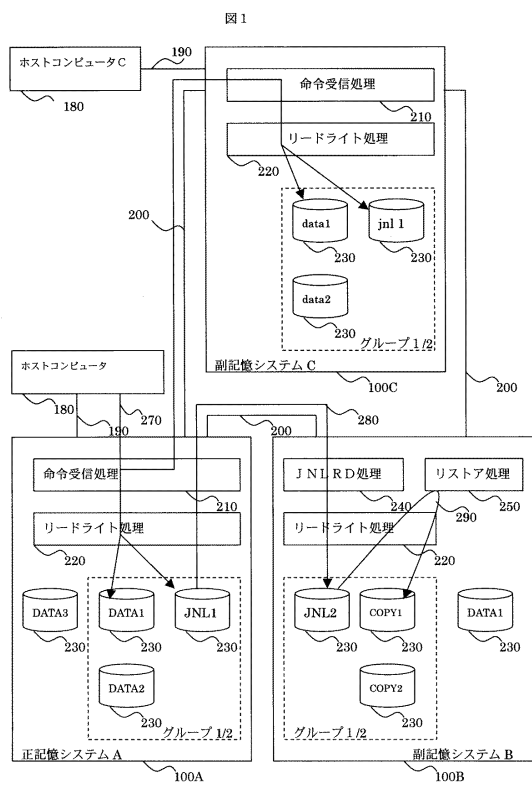
50

## 【 0 2 1 5 】

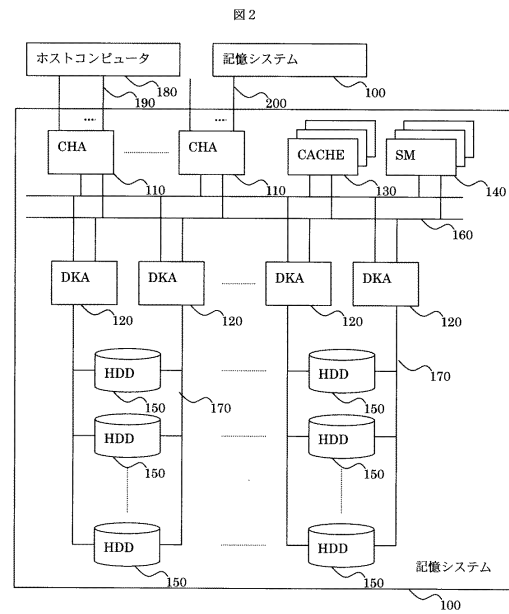
- 1 0 0 記憶システム
- 1 1 0 ホストアダプタ
- 1 2 0 ディスクアダプタ
- 1 3 0 キャッシュメモリ
- 1 4 0 共有メモリ
- 1 5 0 記憶装置
- 1 6 0 コモンバス
- 1 7 0 ディスクアダプタと記憶装置間の接続線
- 1 8 0 ホストコンピュータ

10

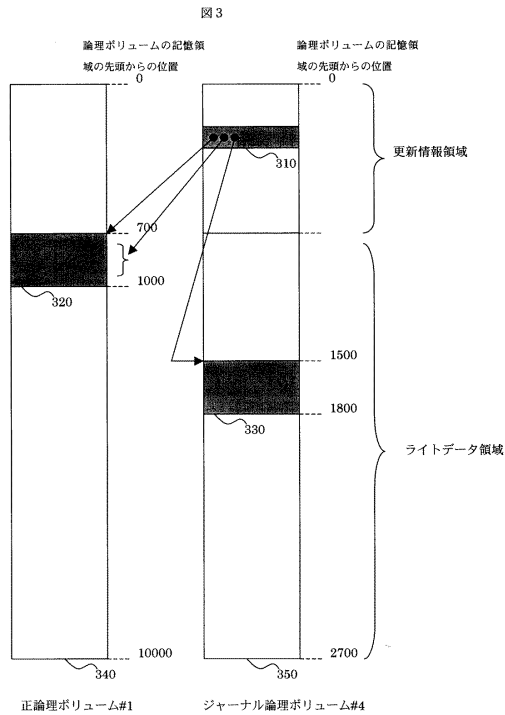
【 図 1 】



【 図 2 】



【図 3】



【図 4】

図 4

論理ボリューム番号	ボリューム状態	フォーマット形式	容量	同期ペア番号	非同期ペア番号	物理アドレス	
						記憶装置番号	先頭から位置
1	正	OPEN3	3	3	1	1	0
2	正	OPEN6	6	4	2	1	3
3	未使用	OPEN6	6	0	0	1	9
4	正常	OPEN9	9	0	0	2	0
5	正常	OPEN3	3	0	0	2	9
6	未使用	OPEN6	6	0	0	2	12

400 ボリューム情報

【図 5】

図 5

ペア番号	ペア状態	正記数システム番号	正論理ボリューム番号	副記数システム番号	副論理ボリューム番号	グループ番号	コピー済みアドレス
1	正常	A	1	B	1	1	0
2	正常	A	2	B	2	1	0
3	正常	A	1	C	1	2	0
4	正常	A	2	C	2	2	0
5	未使用	0	0	0	0	0	0

500 ペア情報

【図 6】

図 6

グループ番号	グループ状態	ペア集合	ジャーナル論理ボリューム番号	更新番号	複製種類	相手記憶システム番号	相手グループ番号	デルタ予約オプション
1	正常	1,2	4	5	非同期	B	1	0
2	正常	3,4	0	0	同期	C	1	0

600 グループ情報

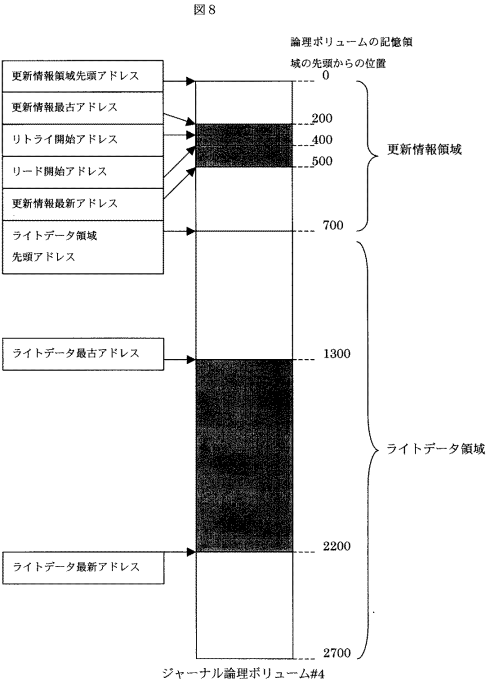
【図 7】

図 7

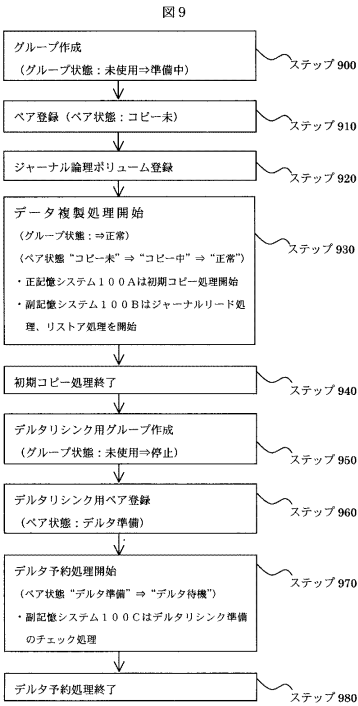
	論理アドレス	
	論理ボリューム番号	論理ボリュームの記憶領域の先頭からの位置
更新情報領域先頭アドレス	4	0
ライトデータ領域先頭アドレス	4	700
更新情報最新アドレス	4	500
更新情報最古アドレス	4	200
ライトデータ最新アドレス	4	2200
ライトデータ最古アドレス	4	1300
リード開始アドレス	4	400
リトライ開始アドレス	4	300

700 ポインタ情報

【図 8】

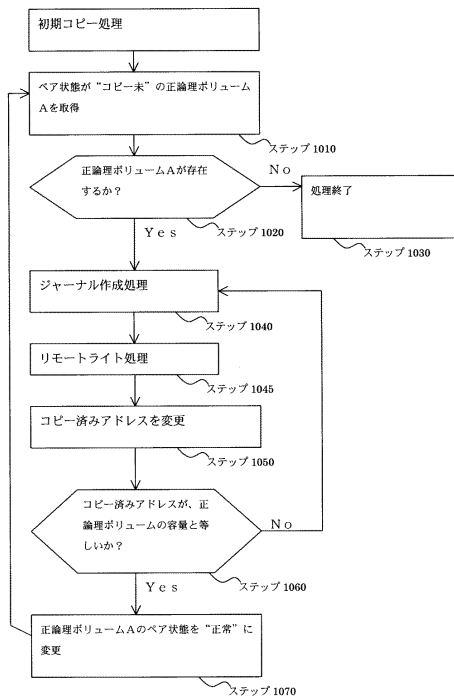


【図 9】



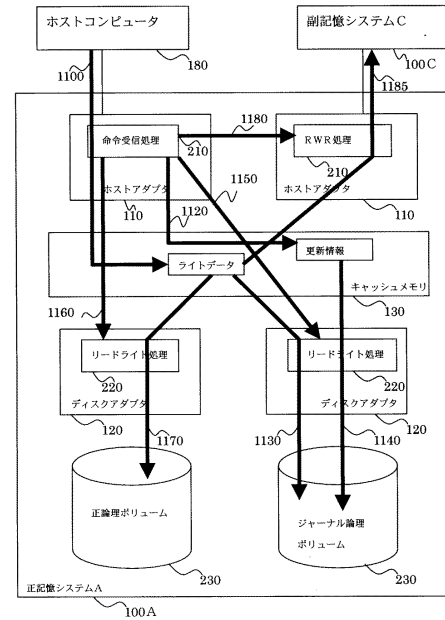
【図 10】

図 10



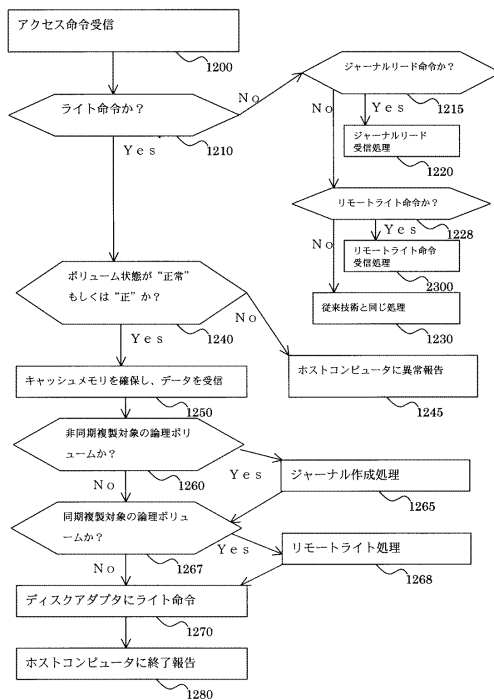
【図 11】

図 11



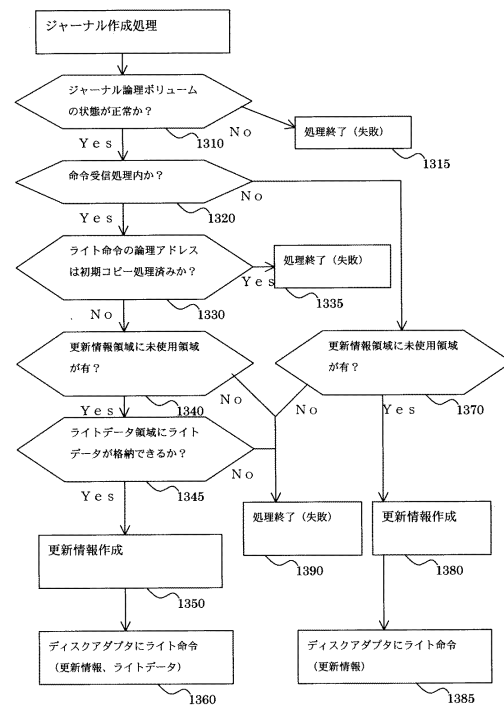
【図 12】

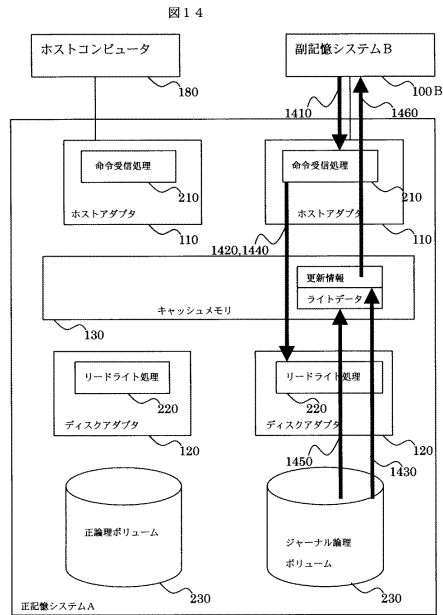
図 12



【図 13】

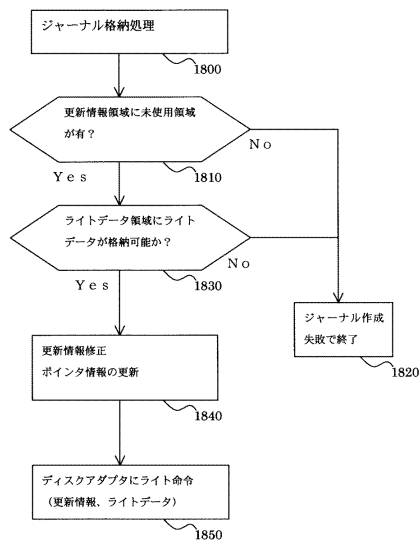
図 13





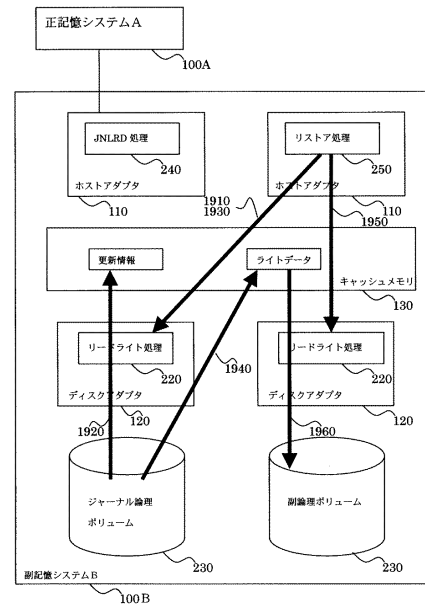
【 図 1 8 】

图 18



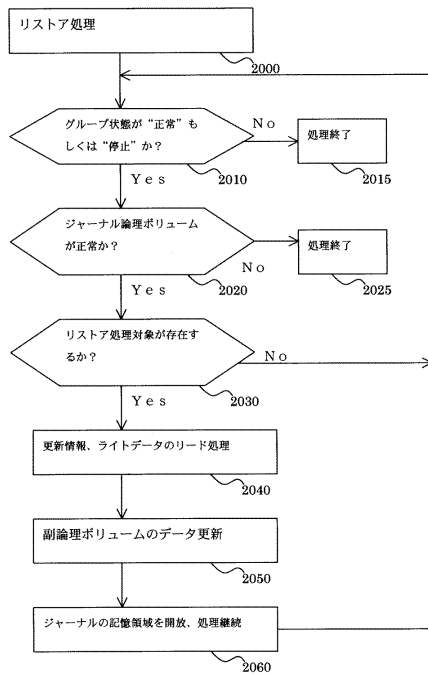
【 図 1 9 】

图 19



【 図 2 0 】

20



【 図 2 1 】

图 2-1

設定項目	設定値例
ライト命令を受信した時刻	1999/3/17 22:20:10
グループ番号	1
更新番号	4
ライト命令の論理アドレス	論理ボリューム番号: 1 論理ボリュームの記憶領域の先頭からの位置: 700
ライトデータのデータサイズ	300
ライトデータを格納したジャーナル論理ボリュームの論理アドレス	論理ボリューム番号: 4 論理ボリュームの記憶領域の先頭からの位置: 1500

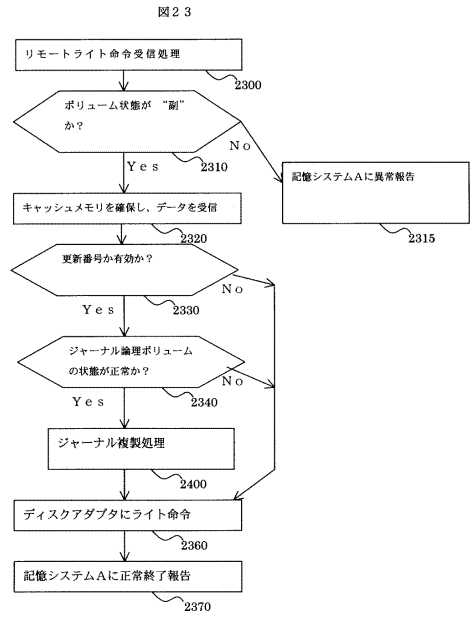
310 更新情報

【 図 2 2 】

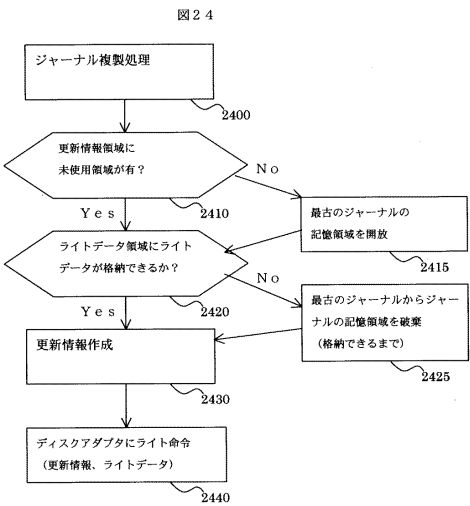
图 2-2

設定項目	設定値例
ライト命令を受信した時刻	1999/3/17 22:22:10
グループ番号	1
更新番号	5
ライト命令の論理アドレス	論理ポートリユーム番号: 1 論理ポートリユームの記憶領域の先頭からの位置: 800
ライトデータのデータサイズ	100
ライトデータを格納したジャーナル論理ポートリユームの論理アドレス	論理ポートリユーム番号: 4 論理ポートリユームの記憶領域の先頭からの位置: 2200

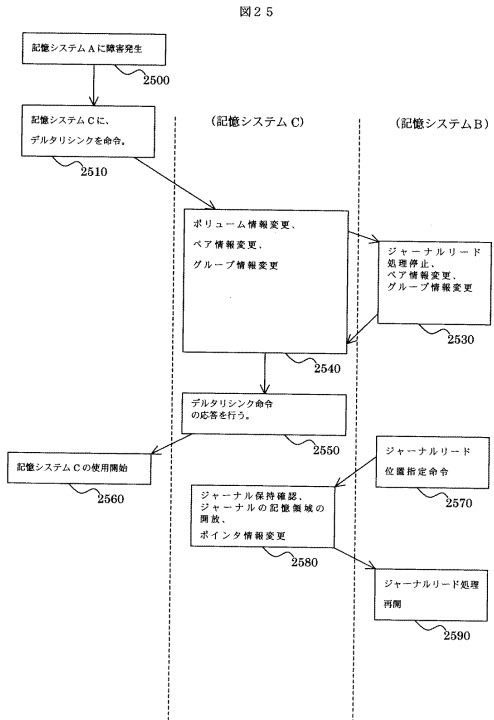
【図 2 3】



【図 2 4】



【図 2 5】



【図 2 6】

図 2 6

論理 ボリューム 番号	ボリューム 状態	フォーマット 形式	容量	同期 ペア番号	非同期 ペア番号	物理アドレス	
						記憶装置 番号	先頭から位置
1	副	OPEN3	3	0	1	1	0
2	副	OPEN6	6	0	2	1	3
3	正常	OPEN9	9	0	0	2	0

【図 2 7】

図 2 7

ペア 番号	ペア 状態	正記憶 システム 番号	正論理 ボリューム 番号	副記憶 システム 番号	副論理 ボリューム 番号	グループ 番号	コピー 済みア ドレス
1	正常	A	1	B	1	1	0
2	正常	A	2	B	2	1	0
3	デルタ待機	C	1	B	1	2	0
4	デルタ待機	C	2	B	2	2	0

【図 2 8】

図 2 8

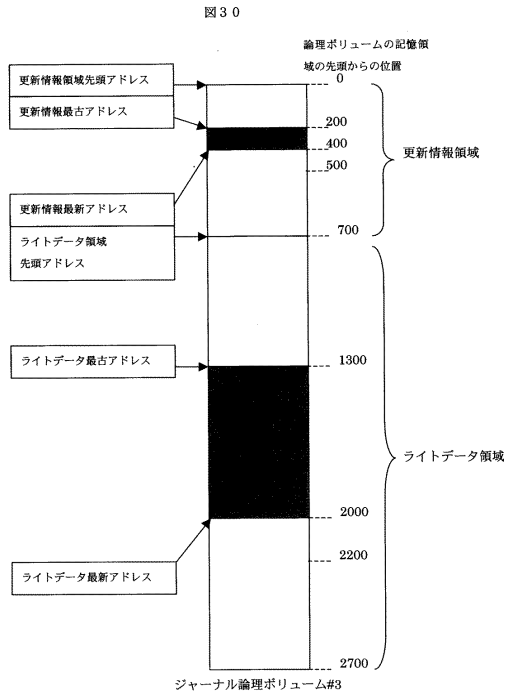
グループ 番号	グループ 状態	ペア集合	ジャーナル 論理ボリ ーム番号	更新 番号	複製種類	相手記憶 システム 番号	相手グループ 番号	デルタ予約 オプション
1	正常	1,2	3	4	非同期	A	1	0
2	停止	3,4	3	0	非同期	C	2	1

【図 2 9】

図 2 9

論理アドレス		
論理ボリューム番号		
論理ボリュームの記憶領域の先頭からの位置		
更新情報領域先頭アドレス	3	0
ライトデータ領域先頭アドレス	3	700
更新情報最新アドレス	3	400
更新情報最古アドレス	3	200
ライトデータ最新アドレス	3	2000
ライトデータ最古アドレス	3	1300
リード開始アドレス	3	0
リトライ開始アドレス	3	0

## 【図 30】

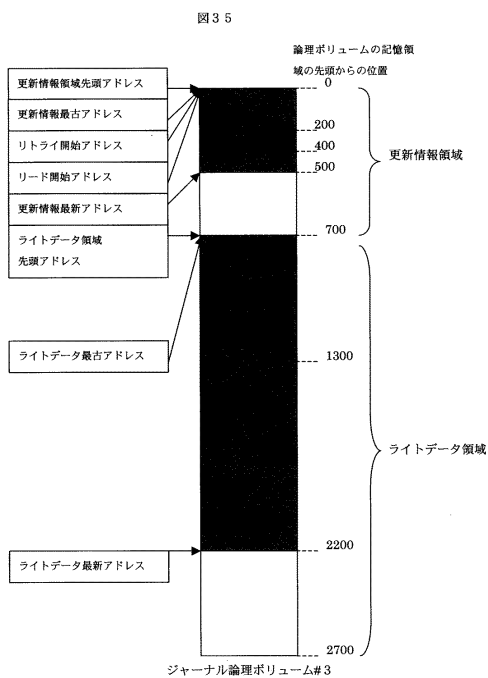


## 【図 31】

図 31

論理ボリューム番号	ボリューム状態	フォーマット形式	容量	同期ペア番号	非同期ペア番号	物理アドレス	
						記憶装置番号	先頭から位置
1	副	OPEN3	3	1	0	1	0
2	副	OPEN6	6	2	0	1	3
3	正常	OPEN9	9	0	0	2	0

## 【図 35】



## 【図 36】

図 36

ペア番号	ペア状態	正記憶システム番号	正論理ボリューム番号	副記憶システム番号	副論理ボリューム番号	グループ番号	コピー済みアドレス
1	デルタ準備	A	1	B	1	1	0
2	デルタ準備	A	2	B	2	1	0
3	正常	C	1	B	1	2	0
4	正常	C	2	B	2	2	0

## 【図 32】

図 32

ペア番号	ペア状態	正記憶システム番号	正論理ボリューム番号	副記憶システム番号	副論理ボリューム番号	グループ番号	コピー済みアドレス
1	正常	A	1	C	1	1	0
2	正常	A	2	C	2	1	0
3	デルタ待機	C	1	B	1	2	0
4	デルタ待機	C	2	B	2	2	0

## 【図 33】

図 33

グループ番号	グループ状態	ペア集合	ジャーナル論理ボリューム番号	更新番号	複製種類	相手記憶システム番号	相手グループ番号	デルタ予約オプション
1	正常	1,2	3	5	同期	A	2	0
2	停止	3,4	3	1	非同期	B	2	1

## 【図 34】

図 34

	論理アドレス	論理ボリューム番号	論理ボリュームの記憶領域の先頭からの位置
更新情報領域先頭アドレス	3	0	0
ライトデータ領域先頭アドレス	3	700	700
更新情報最新アドレス	3	500	500
更新情報最古アドレス	3	0	0
ライトデータ最新アドレス	3	2200	2200
ライトデータ最古アドレス	3	700	700
リード開始アドレス	3	0	0
リトライ開始アドレス	3	0	0

## 【図 37】

図 37

グループ番号	グループ状態	ペア集合	ジャーナル論理ボリューム番号	更新番号	複製種類	相手記憶システム番号	相手グループ番号	デルタ予約オプション
1	停止	1,2	3	4	非同期	A	1	1
2	準備中	3,4	3	4	非同期	C	2	0

## 【図 38】

図 38

論理ボリューム番号	ボリューム状態	フォーマット形式	容量	同期ペア番号	非同期ペア番号	物理アドレス	
						記憶装置番号	先頭から位置
1	正	OPEN3	3	1	3	1	0
2	正	OPEN6	6	2	4	1	3
3	正常	OPEN9	9	0	0	2	0

## 【図 39】

図 39

ペア番号	ペア状態	正記憶システム番号	正論理ボリューム番号	副記憶システム番号	副論理ボリューム番号	グループ番号	コピー済みアドレス
1	異常	A	1	C	1	1	0
2	異常	A	2	C	2	1	0
3	正常	C	1	B	1	2	0
4	正常	C	2	B	2	2	0

## 【図 40】

図 40

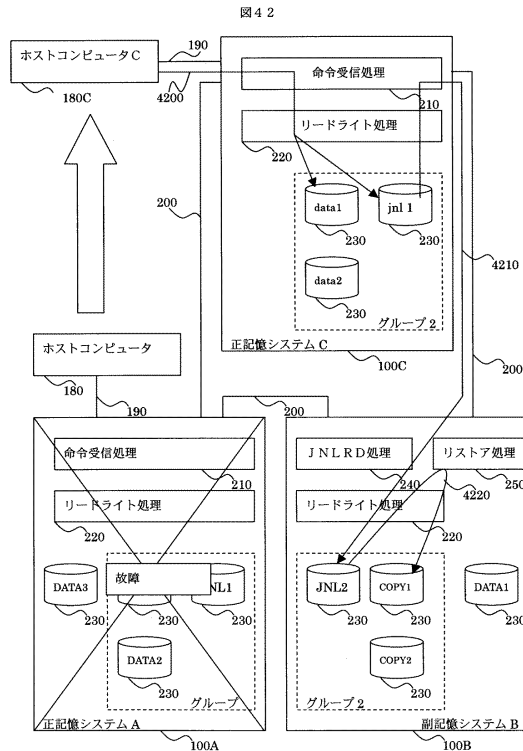
グループ番号	グループ状態	ペア集合	ジャーナル論理ボリューム番号	更新番号	複製種類	相手記憶システム番号	相手グループ番号	デルタ予約オプション
1	異常	1,2	3	5	同期	A	2	0
2	正常	3,4	3	5	非同期	B	2	0

## 【図 41】

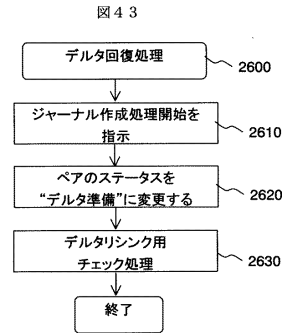
図 41

	論理アドレス	論理ボリューム番号	論理ボリュームの記憶領域の先頭からの位置
更新情報領域先頭アドレス	3	0	0
ライトデータ領域先頭アドレス	3	700	700
更新情報最新アドレス	3	500	500
更新情報最古アドレス	3	400	400
ライトデータ最新アドレス	3	2200	2200
ライトデータ最古アドレス	3	2000	2000
リード開始アドレス	3	400	400
リトライ開始アドレス	3	400	400

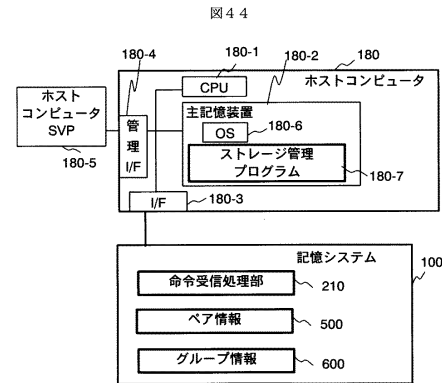
【図 4 2】



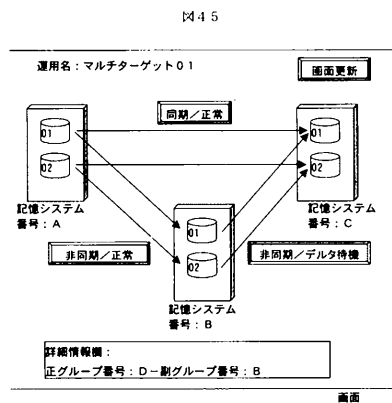
【図 4 3】



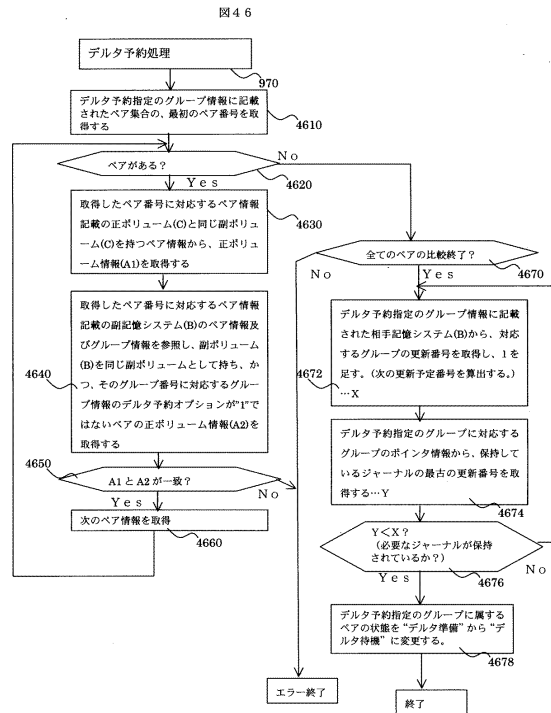
【図 4 4】



【図 4 5】



【図 4 6】



---

フロントページの続き

(72)発明者 今津 剛行

神奈川県横浜市戸塚区戸塚町5 0 3 0 番地 株式会社日立製作所 ソフトウェア事業部内

審査官 桜井 茂行

(56)参考文献 特開2 0 0 5 - 0 8 4 9 5 3 ( J P , A )

特開平1 0 - 3 1 2 3 2 7 ( J P , A )

特開2 0 0 3 - 1 2 2 5 0 9 ( J P , A )

(58)調査した分野(Int.Cl. , D B 名)

G 0 6 F 1 2 / 0 0

G 0 6 F 3 / 0 6