

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第4325836号
(P4325836)

(45) 発行日 平成21年9月2日 (2009.9.2)

(24) 登録日 平成21年6月19日 (2009.6.19)

(51) Int. Cl.

F I

G O 6 F 3/06 (2006.01)

G O 6 F 12/00 (2006.01)

G O 6 F 13/10 (2006.01)

G O 6 F 15/00 (2006.01)

G O 6 F 3/06 3 O 4 F

G O 6 F 3/06 3 O 5 F

G O 6 F 12/00 5 1 4 E

G O 6 F 12/00 5 3 1 D

G O 6 F 12/00 5 4 5 A

請求項の数 8 外国語出願 (全 20 頁) 最終頁に続く

(21) 出願番号	特願2002-369171 (P2002-369171)	(73) 特許権者	000005108
(22) 出願日	平成14年12月20日 (2002.12.20)		株式会社日立製作所
(65) 公開番号	特開2003-263280 (P2003-263280A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成15年9月19日 (2003.9.19)	(74) 代理人	100093861
審査請求日	平成17年12月9日 (2005.12.9)		弁理士 大賀 真司
(31) 優先権主張番号	10/042, 376	(72) 発明者	渡辺 直企
(32) 優先日	平成14年1月3日 (2002.1.3)		アメリカ合衆国カリフォルニア州サニーベ
(33) 優先権主張国	米国 (US)		イル アパートメント# 3 2 1 イースト
			エルカミノリアル965
		審査官	木村 貴俊

最終頁に続く

(54) 【発明の名称】 複数リモートストレージのデータ同期方式

(57) 【特許請求の範囲】

【請求項 1】

ローカルストレージシステムと複数のリモートストレージシステムで構成されるシステムにおいて、該複数のリモートストレージシステムに記憶された該ローカルストレージシステムのバックアップデータを同期させるための方法であって、

該各リモートストレージシステムが、該ローカルストレージシステムからのデータ更新要求を受信し、該受信したデータ更新要求の識別子を記録するステップと、

該各リモートストレージシステムが、該ローカルストレージシステムからのデータ更新要求を受信し、該データ更新要求に付随するデータをストレージメディアへ書き込んだデータ更新要求の識別子を記録するステップと、

該リモートストレージシステムのうちのいずれかのリモートストレージシステムが、該ローカルストレージシステムで障害が検出されると、該受信したデータ更新要求の識別子及び該受信し書き込みを完了したデータ更新要求の識別子の該各リモートストレージシステム間での差異を検出するステップと、

該差異を検出したリモートストレージシステムが、該差異に基づいて、該各リモートストレージシステム間のデータを同期させるステップとから構成されるデータ同期方法。

【請求項 2】

処理システムであって、

データ記憶のためのストレージメディアを有するローカルストレージシステムと、

各々付随データを有するリモートコピーコマンドに応答して、データを記憶するための

ストレージメディアを有し、受信したリモートコピーコマンドを識別する情報及び付随データをストレージメディアに書き込んだ、受信したリモートコピーコマンドを識別するための情報を維持する、少なくとも二つのリモートストレージシステムと、
を備え、

該ローカルストレージシステムは、該各リモートストレージシステム内のストレージメディアにデータを書き込ませるために、該各リモートストレージシステムに該リモートコピーコマンドを送信し、

該ローカルストレージシステムでの障害検出により、該リモートストレージシステムのうちの選択されたりリモートストレージシステムは、該各リモートコピーシステムが維持している該受信したリモートコマンドを識別するための情報を共有して、一つのリモートストレージシステムでは受信未了で、他のリモートストレージシステムでは受信完了しているリモートコピーコマンドを特定して、該各リモートストレージシステム間のデータを同期させる、処理システム。

10

【請求項 3】

該各リモートストレージシステムは、リモートコピーコマンドと共に受信されたデータを送受信する手段を有する、請求項 2 に記載の処理システム。

【請求項 4】

該各リモートストレージシステムは、一つのリモートストレージシステムで受信済みのデータを他のリモートストレージシステムに送信する為に、通信手段を用いる、請求項 3 に記載の処理システム。

20

【請求項 5】

処理システムであって、

データ記憶のためのストレージメディアを有するローカルストレージシステムと、

複数のリモートストレージシステムであって、それぞれがリモートコピーコマンドにตอบสนองして受信したデータを保存するストレージメディアを有する複数のリモートストレージシステムと、

を備え、

該ローカルストレージシステムは、該各リモートストレージシステム内のストレージメディアにデータを書き込ませるために、該各リモートストレージシステムにリモートコピーコマンドを送信し、

30

該ローカルストレージシステムでの障害検出により、該リモートストレージシステムのうちの選択されたりリモートストレージシステムは、該各リモートストレージシステムが保有する更新情報の受信の差異に関する情報を共有して、一つのリモートストレージシステムで受信されている情報を、他の未受信のリモートストレージシステムに送信して、リモートストレージシステム間で、互いにデータがミラーとなるように修正する、処理システム。

【請求項 6】

各リモートストレージシステムは、該各リモートコピーシステムが維持している該リモートコピーコマンドを識別するための情報を保持するキューを含む、請求項 5 に記載の処理システム。

40

【請求項 7】

該情報は、受信されたりリモートコピーコマンドの記録を有する、請求項 6 に記載の処理システム。

【請求項 8】

該情報は、該リモートストレージシステムのストレージメディアに書き込み済みの受信リモートコピーコマンドの記録を有する、請求項 7 に記載の処理システム。

【発明の詳細な説明】

【発明の属する技術分野】

本発明は、一般的に、ローカルストレージシステムと、このローカルストレージシステムが保持するデータの少なくとも一部をミラーする複数のリモートストレージシステムとか

50

ら構成されるデータストレージ処理システムに係る。更に具体的には、本発明は、これらのシステム中の一システムが障害になった場合に、生存するストレージシステムのデータを同期させるための、方法及びこの方法を実装した装置に関連する。

【従来の技術】

商業、軍、政府、やその他の事業での積年のデータ処理の利用により大量のデータが保存されており、この多くのものは、実質上金銭に換算できないほどの重要性を持っている。例えばデータが失われると、企業は倒産する危険がある。偏にこのために、ローカルデータは一つ以上のコピーでバックアップされており、元データが万一失われても利用できるように保持されている。データが重要であるほど、バックアップ方法は精密化されなければならない。

10

重要データを保護する一つの方法は、例えば、ローカルストレージシステムから地理的に離れた一つ以上のサイトで、そのデータのバックアップコピーを保存することである。各リモートストレージシステムがローカルストレージシステムのデータのミラーイメージを保持し、ローカルストレージシステムのローカルデータイメージに変更(書き込み、削除等による)が生じた場合には、該変更データイメージを各リモートストレージシステムに転送して、各リモートストレージシステムがローカルデータイメージのミラーを維持できるようにする。ローカルストレージシステムのデータをリモートストレージシステムがミラーする一つの例が米国特許 No. 5,933,653 に開示されている。

リモートストレージシステムに送信される更新データはしばしばキューされ、リモートコピー動作のオーバーヘッドを最小にするために、グループとして送信される。又転送媒体には、インターネットまたは類似媒体がよく使用される。このような理由で、ローカルデータをミラーしているデータイメージと言えども、各時点では必ずしも同じではない。もし、ローカルデータのミラーのために複数のリモートストレージが使用されている場合には、各リモートストレージ間でデータイメージが一致していないことがしばしば発生する。少なくとも、ローカルストレージシステムによるデータを更新し終わる迄の時間は不一致になってしまう。このような異なったデータイメージが存在する事は、ローカルストレージシステムが障害になり、リモートストレージシステムが取り残された時には、問題になる。

20

ローカルストレージシステムが障害になった時点では、あるリモートストレージは、完全なミラーとまではいなくても、最新に最も近いローカルストレージシステムのデータイメージを保持しており、他のリモートストレージでは、ある種のデータイメージは全く更新されていないという事態が発生する。かくのごとく、ローカルストレージシステムが障害になった場合に、全システムを再起動させるには、他のリモートストレージシステムのデータを同期させて、各システムが最新に最も近い同じデータイメージを保持することが必要になる。幾つかのデータ同期方式が存在する。

30

もし、ローカル及びリモートストレージシステムで可搬媒体(例えば、テープ、CD-R、DVD等)が使用出来る場合は、そのような可搬媒体を用いることが出来る。例えば、システム管理者は、ローカルストレージシステムの最新に最も近いデータイメージを保持していると信じられるリモートストレージシステムのデータをテープにコピーする。

次いで、他のリモートストレージシステムで同期させるために使われる前にデータイメージが変更されてしまうのを防ぐために、このテープを持ちまわって他のリモートストレージシステムのデータを更新し終えるまで、データイメージを供給するストレージシステムの入力/出力(I/O)動作は停止される。リモートストレージシステムでは、管理者が可搬媒体からリモートサイトのストレージにコピーする。コピーが終わった所で、システム管理者はこれまではリモートストレージシステムの一つであったものが新たなローカルストレージシステムになるように全システムを指定替えをして、I/O動作が開始できることになる。この方式は、関連するデータ量が少なく、それほど大きくないシステムに有効である。より大きなシステムではデータ量が急速に増大し、全同期プロセスに必要なコピー時間が過大になる。

40

可搬媒体が使用できない場合は、別の方式として、いろいろなストレージシステム間のデ

50

ータ転送にネットワーク接続を使用する方式が考えられる。この方式では、一つのストレージシステムが、前の、(現在は障害になった)ローカルストレージシステムの代わりを務めるために選択される。

この選択されたストレージシステムは、先に述べたと同じ理由で、I/O動作は停止され、この選択されたストレージシステムと他のリモートストレージシステムとの間で、同期コピー処理が開始される。同期処理が完了すると、選択されたストレージシステムにおいてI/O動作が再スタートされ、(以前の、障害になったローカルストレージシステムの分だけ)ストレージシステムが一つ少なくともはなるがシステムは以前と同様に業務を続けることができる。

この後者の方式の大きな問題は、同期処理に必要な時間であり、特に大量データを保有する大システムでは問題が大きい。例えば、100TBのデータを100MB/sの転送速度のネットワークで同期化する場合の時間は11.57日が必要となる。 $(100 \times 10^{12} / (100 \times 10^6)) = 10^6 \text{秒} = 277 \text{時間} = 11.57 \text{日}$ 。この時間はただ一つのストレージシステムを同期するのに必要な時間である。同期を必要とするストレージシステムが増加すると、この問題は深刻化する。更に、同期処理中は、関係するストレージシステムのI/O動作は停止しなければならない。

【発明が解決しようとする課題】

本発明は、ローカルストレージの障害時に於いても、複数のリモートデータストレージシステムが同じデータイメージを保持できるように、当該リモートデータストレージシステムを同期させるための方法及び当方法を実装するための方式を提供する。

【課題を解決するための手段】

一般的に本発明は、複数のリモートストレージシステムに通信結合したローカルデータストレージシステムから成る、データ処理システムに関係する。各ストレージシステムはローカルであるにせよ、リモートであるにせよ、データストレージにストレージメディアを保有する。ローカルデータストレージシステムのストレージメディアにて維持されているデータは、各リモートストレージシステム内のストレージメディアにミラーされる。

ローカルストレージシステムで発生したデータイメージの変更は、更新データと共にデータメッセージを送信する、リモートコピープロセスを用いて定期的にもリモートストレージシステムに送信され、当該システムのデータイメージの更新が行われる。各ストレージシステムは、リモートストレージシステムがどんな更新データを受信したかのヒストリを示す情報、及びどんな更新データを(かかるリモートストレージシステムの各ストレージメディアへ書き込むことで)取り込んだかのヒストリを示す情報を保持している。一つのストレージシステムが障害になった場合は、生存ストレージシステムが、このヒストリの更新データを持ち回って、システム間のデータイメージの差異、即ち、各生存ストレージシステムでの未受領の更新データの有無を調べる。もし差異が存在する場合には、生存ストレージシステムがデータイメージの同期を実施して、全ての生存システムが実質的に同一データイメージを保有するようにする。

本発明の一実施例では、同期は、ヒストリの更新データ情報が示す最新の更新データを有するリモートストレージシステムが、他のリモートストレージシステムに必要な更新データを送信して、全てのデータイメージを最新状態に持ち上げる“ロールフォワード”方式で達成される。もう一つの“ロールバック”方式の同期方式では、不揃いの更新データを切り捨てて、すべてのデータイメージを共通の更新レベルに戻してしまう。

本発明の一つの効果は、ミラーされるストレージを有するデータ処理システムに於いては、ミラーするローカルストレージシステムに障害が起こった場合、ローカルストレージのミラーイメージは、互いに、対応付けが出来るという事実にある。

本発明のもう一つの実施例では、各ストレージシステムはキュー構造を持ち、ロールバックキューでは、ストレージメディアに未書き込みのメッセージを識別し、ライトヒストリキューでは、実際に書き込みが完了したメッセージを識別する。もし、ローカルストレージシステムで障害が発生したら、各リモートストレージシステムは、各システムのロールバックキューとライトヒストリキューの内容を示した情報を持ち回って、一つのストレ

10

20

30

40

50

ジシステムでは保有しているが他のストレージシステムでは保有していないデータを見つけ出し、この情報に基づいて不揃いデータを、保有しているシステムから保有していないシステムに送信する。

本発明の更に他の実施例では、アラインメント処理を有し、ローカルストレージシステムがリモートストレージシステムで受信されたメッセージの来歴を追いかけて、当該情報により、各ロールバック及びライトヒストリキュー内に存在するメッセージを特定する。後になって、ロールバック、ロールフォワード、又はパージ処理が実施できるように、各ロールバックキュー及びライトヒストリキューに登録されているメッセージのオーバーラップを維持するために、ローカルストレージシステムはデータ転送を抑制し調整する。

本発明の、これら及び他の効果が、添付する図面と共にこの明細書の以降に記載する詳細な説明により、本分野に造詣のある人には明確になる。

【発明の実施の形態】

図1は、データ通信ネットワーク16で互いに通信結合される、ローカルサイト12と多数のリモートサイト14を含む、データ処理システム10を説明する。ローカルサイト12は、ホストプロセッサ101とローカルストレージシステム104を含む。ホストプロセッサ101とローカルストレージシステム104は、ローカルサイト12でのデータイメージの更新データをリモートサイト14と交信するために、データ通信ネットワーク16を介してリモートサイト14に結合している。これにより、リモートサイト14は、ローカルサイト12のデータのミラーとなるデータイメージを維持することが出来る。

各リモートサイト14は、リモートストレージシステム105, 106を有し、対応するホストプロセッサ102, 103を持つ場合と持たない場合がある。リモートストレージシステム105, 106は、ローカルストレージシステム104のデータイメージのミラーとなるデータイメージを保持するために、ローカルストレージシステム104のストレージメディアと極めて類似したストレージメディアを有する(後に更に議論する)。

リモートサイトは、ローカルストレージシステム104が計画的又は予期しない停止になった場合には、必要データへのアクセスを可能にするが、ローカルサイト12が災害でデータが失われる場合のことを考えると、リモートサイト14は、サイト及び維持するデータを保護するために、ローカルストレージシステム104から地理的に離れた位置に設置することがより大きな効果をもたらす。

リモートストレージシステム105, 106は、ローカルストレージシステム104と実質的には、同一であり、ローカルストレージシステムに対する議論は、等しくリモートストレージシステム105, 106にも適用される。

ローカルサイト12では、ホストプロセッサ101は、ネットワークインターフェース(I/F)111とネットワーク結合111aを通して、ネットワーク16に結合し、入力/出力バス(I/Oバス)108とI/Oインターフェース110を通して、ローカルストレージシステム104に結合する。

ローカルストレージシステム104は、ディスクコントローラ141を有し、このディスクコントローラは、I/Oバス108を通してホストからのI/O要求を受け取るための、I/Oインターフェース130を持ち、ネットワークインターフェース(I/F)131とネットワーク結合131aを通して、データ通信ネットワーク16に接続される。更に、このローカルストレージシステム104は、データバス132とI/Oインターフェース137, 138を通してこのディスクコントローラ141に接続される、ディスクユニット140で構成されるストレージメディア142を有する。

このディスクコントローラ141自身、内部バス132を通して、メモリ134に結合するCPU(Central Processing Unit: 中央処理装置)133を持ち、この内部バス132は、ディスクコントローラの多様なインターフェース(例えば、I/O I/F 130, 138, ネットワークI/F 131等)にも接続される。メモリ134はキャッシュメモリ135を含み、このキャッシュメモリはホストプロセッサ101からのI/O要求に応じて、ストレージメディア142との間の読み書き時のキャッシュ用として利用される。このメモリは、制御情報136のようなある種のデータ構造や情

10

20

30

40

50

報を保持するために使用される。

C P U 1 3 3 は、ストレージメディア 1 4 2 のデータイメージへの（例えば、ホストプロセッサ 1 0 1 から要求された）変更データを、通常のやり方で（自らのストレージメディアを持つ）リモートストレージシステム 1 0 5 , 1 0 6 に送信処理する、リモートコピープロセスを優先的に実行する。これにより、リモートストレージシステム 1 0 5 , 1 0 6 は、ローカルストレージシステム 1 0 4 のデータをミラーできる。簡単に言うと、リモートコピープロセスは以下の通り処理される：ディスクコントローラ 1 4 1 は、ストレージメディア 1 4 2 上に保持されているデータ（データイメージ）への追加、変更、削除、その他の修正のいずれかを行う書き込み I / O 要求を受け取ると、当該ストレージメディア 1 4 2 へデータの書き込みを行う。そのようなストレージメディアへの書き込みまたは、少なくとも書き込みの一部でも、リモートサイト 1 4 でミラーされたデータイメージが必要な場合には、データメッセージが生成されて、リモートコピーキュー（図示されていない）に登録される。このリモートコピーキューは、C P U 1 3 3 上で走行している、リモートコピープロセスにより周期的にチェックされる。もし、このキュー内に一つ以上のデータメッセージが検出されたら、このメッセージは抽出されて、各リモートストレージシステム 1 0 5 , 1 0 6 に送信され、データメッセージのデータが書き込まれ、それによって、それらのデータイメージを更新してローカルサイト 1 2 のデータイメージに一致させる。データメッセージに関する情報は、ストレージシステム 1 0 4 , 1 0 5 , 1 0 6 にて、ローカルストレージシステム 1 0 4 から送信された、又は送信途中のデータメッセージか、又は、リモートストレージシステムで受信されたデータメッセージか、を識別するヒストリ情報として保持される。各ストレージシステムでは、ヒストリ情報を複数のキュー構造により実装する。図 2 はこれらのキュー構造を示す。

図 2 に示す通り、ローカルストレージシステム 1 0 4 は、ワーキングキュー 1 1 0、ロールバックキュー 1 1 1、中間キュー 1 1 2、及びライトヒストリキュー 1 1 3 の各キューを含む。ワーキングキュー、ロールバックキュー、中間キュー、及びライトヒストリキュー（1 1 0 - 1 1 3）は、リモートストレージシステム 1 0 5（キュー 1 2 0, 1 2 1, 1 2 2、及び 1 2 3）及びリモートストレージシステム 1 0 6（キュー 1 3 0, 1 3 1, 1 3 2、及び 1 3 3）にそれぞれミラーされる。各キューは、F I F O（F i r s t - I n - F i r s t - O u t：先入れ先出し）構造で実装される。

典型的な読み書き I / O 要求は、コマンドエントリとこれに伴うか又は後続するデータ（書き込み要求の場合）を含む。各コマンドエントリには、データと、データが読み書きされるストレージメディア 1 4 2 のロケーション及び、実装に従って必要とするその他の情報が指定される。もしホストプロセッサ 1 0 1 から、リモートストレージシステム 1 0 5 , 1 0 6 でミラーされるデータイメージを変更するデータを書き込むための I / O 要求を受けたら、該コマンドエントリには順序番号がアサインされる。するとコマンドエントリはコマンドエントリとアサインされた順序番号により形成されたものになる。このコマンドエントリは次に、ワーキングキュー 1 1 0 に登録される。これは、データに先立って、リモートストレージシステムに送信される、データとデータメッセージのためのヒストリ情報を形成する。

コマンドエントリがワーキングキュー 1 1 0 にある間に、対応する書き込み要求は受信処理されている。この処理は、書き込み要求に対応するデータ（ホスト 1 0 1 とストレージシステム 1 0 4 に適用されている通信プロトコル次第で、要求と共に転送されるか、又は要求の後に続いて転送されるかどうかである）の受信、キャッシュ 1 3 5 にデータを格納するためのキャッシュエリアのアサイン、及び受信時の、このエリアへのデータの書き込み、を含む。キャッシュされたデータへのポインタは、対応するコマンドエントリに関連付けられる。この I / O 要求の送信者に対して、状態メッセージが返送されて、この I / O 要求のデータ状態即ち受信できたか、又は受信時エラーが検出されたかが報告される。

図 2 は、ディスクコントローラ 1 0 4 で受信され、順序番号 “ 1 5 ” がアサインされた書き込み I / O 要求 1 0 1 a と、アサインされた順序番号と書き込み要求のコマンドエント

10

20

30

40

50

リで形成されたコマンドエントリを示す。コマンドエントリは既に述べられたように更に処理する為に、ワーキングキュー 110 に登録される。

当該 I/O 要求が処理され、エラーが検出されない場合、このコマンドエントリは、ワーキングキュー 110 からロールバックキュー 111 に移される。ロールバックキューは一時的な保持エリアで、後により詳しく説明される、ローカルストレージシステム 104 とリモートストレージシステム 105, 106 の間のロールバック同期のために使用される。リモートストレージシステム 105, 106 は、ロールバックキュー含めて、基本的には、同じ目的のために、同じキュー構造を持っている。ここでは、ロールバックキューに含まれているコマンドエントリに対応するデータは、本発明では、同期プロセスで使用され、ストレージシステムの一つが、障害になった場合には、破棄されるか、生存ストレージシステム間で比較するために使用される、とだけ言うておけば十分であろう。

コマンドエントリは、ロールバックキュー 111 から、ストレージメディア 142 への書き込み待ちの間又は、書き込んでいる間エントリを保持する中間キュー 112 に移される。当該ストレージメディアへの書き込みが終了すると、当該コマンドエントリは、プライマリ（ローカル）ストレージシステム 104 のミラーコピーを行うリモートストレージシステムに対してデータイメージの更新を行わせるために送信される、データメッセージを作成してリモートコピー要求を作成するために利用される。次いで、キャッシュポインタはコマンドエントリから削除され、エントリは中間キューからライトヒストリキュー 113 に移される。図 2 は、既に示されたように、ホストプロセッサ 101 から受信した順序番号 15 をアサインされ、ワーキングキュー 110 に登録されているコマンドエントリを示す。現在ロールバックキュー 111 に登録されている順序番号 13, 14 のコマンドエントリは、関連データが、ストレージメディア 142 に書き込まれるのを待っている。中間キュー 112 は、順序番号 10, 11、及び 12 のコマンドエントリを保有している。順序番号 10 のコマンドエントリは、次に書き込まれるか、書き込まれている最中であることを示している。順序番号 7, 8, 及び 9 のコマンドエントリは、関連データをストレージメディア 142 に書き込み済で、それ故にライトヒストリキュー 113 に登録されている。

リモートストレージシステムは基本的には、同じキュー構造を持ち、同様に操作される。例えば、リモートストレージシステム 105 は、当該エントリの全データパッケージを受信している間、ワーキングキュー 120 に登録されている順序番号 10 がアサインされたコマンドエントリを持つ、データメッセージを受信中であることを示している。この順序番号 10 はローカルストレージシステム 104 で与えられたものである。一度、データメッセージが、受信されると、そのコマンドエントリは、ワーキングキュー 120 からロールバックキュー 121 に移される。このロールバックキューには現在、順序番号 6 - 9 を持つデータメッセージのコマンドエントリが登録されている。

中間キュー 122 には順序番号 5 のデータメッセージのコマンドエントリが存在し、ストレージシステム 105 のストレージメディア 142' への書き込み中であることを示している。書き込みが終了すると、当該エントリは、順序番号 1 - 4 のデータメッセージのコマンドエントリが既に存在するライトヒストリキュー 123 に移動する。ライトヒストリキュー 123 の深さにより、最古のエントリ、例えば順序番号 1 のエントリは順序番号 5 のコマンドエントリの加入により追い出されることになる。

他のリモートストレージ 106 も同じデータキュー構造 (130, 131、・・・133) を持つ。図 2 では、リモートストレージ 106 が、現在順序番号 12 のデータメッセージを受信しつつあり、ワーキングキュー 130 に順序番号 12 のコマンドエントリが入っている。ロールバックキュー 131 には、順序番号 8 - 11 を持つ制御情報が登録されている。以下同様である。

このようなヒストリ情報をストレージシステムによりトレースするキューは通常、メモリに保持されているが、ストレージメディアに保持してもかまわないことは、本分野に造詣のある人には明らかである。

ローカルストレージシステム 104 は、更に、メモリ 134 (図 1) 中にリモートコピー状

10

20

30

40

50

態テーブル 1 1 4 を持ち、どのデータメッセージが送信され、どのリモートストレージシステムによりどのデータメッセージが受信され、完了報告が出されたかを表示している。例えば、最も最近に、リモートストレージシステム 1 0 5 (表 1 1 4 で “ S 1 ” と表示) に受け取られたのは、順序番号 9 であり、リモートストレージシステム 1 0 6 (“ S 2 ” と表示) に受け取られたデータメッセージは、順序番号 1 1 であることが分かる。

リモートコピー状態テーブル 1 1 4 は更に、リモートストレージシステムのロールバックキュー、ライトヒストリキューに関する情報も保持している。即ち、テーブル 1 1 4 が示すようにリモートストレージシステム 1 0 5 , 1 0 6 のロールバックキューの長さはデータメッセージ 4 エントリ分で、1 0 M B の容量まで収容可能であることが分かる。各リモートストレージシステム 1 0 5 , 1 0 6 のライトヒストリキューは 5 データメッセージの 5 エントリ分である。このライトヒストリキューのバイト単位の容量をこのテーブル 1 1 4 に保持させても良い。

10

図 3 は各種のネットワークアーキテクチャを持つデータ処理システムで発生し得る、種々の障害モードのシナリオを示す。最初の図 3 A では、ローカルストレージシステム 3 0 4 A を持つデータ処理システムは、リモートストレージシステム 3 0 5 A , 3 0 6 A に並列に通信接続されている。図 3 A に示すように、リモートストレージシステム 3 0 5 A が障害になった場合は、データ同期プロセスの必要性はない。ローカルストレージシステム 3 0 4 A は依然稼働中であり、生存リモートストレージシステム 3 0 6 A が通常のリモートコピー処理によりデータ更新を継続できるからである。

図 3 B はローカルストレージシステムが障害になった場合のシナリオを図で示す。この場合も、ローカルストレージシステム 3 0 4 B は、リモートストレージシステム 3 0 5 B , 3 0 6 B に並列に通信接続されている。もし、ローカルストレージシステム 3 0 4 B の障害が検出されると、図 3 B の下半分に記されているように、リモートストレージシステム 3 0 5 B , 3 0 6 B は、最新の更新データを保有しているリモートストレージシステムのデータイメージに対して、自己の持つデータイメージを同期させる。リモートストレージシステム 3 0 5 B が最新更新情報を保有しているとする、本発明での同期プロセスは、リモートストレージシステム中のデータイメージの一つを、他システム中のそれと同期することにより行われる。すぐ分かるように同期には二つの方式が存在する。

20

ロールフォワード方式では、最新の更新データがこれを必要としているリモートストレージシステムに送信され、ロールバック方式では、最新の更新データは破棄される。いずれの方式でも、生存リモートストレージシステムが保持しているデータイメージは同一化される。

30

図 3 C , 3 D , 3 E では、ローカル及びリモートストレージシステム 3 0 4 C , 3 0 5 C , 3 0 6 C が各々、デジチエンネットワーク結合で接続されている。図 3 C では、ローカルストレージシステム 3 0 4 C が障害になり、リモートストレージシステム 3 0 5 C がリモートストレージシステム 3 0 6 C を同期させる場合を示している。

一方に於いて、図 3 D の如く、デジチエンの遠端に存在するリモートストレージシステム 3 0 6 C が障害になった場合は、データ同期の必要性はない。図 3 E の如く、中間のリモートストレージシステムが障害になった場合は、残ったリモートストレージシステムは、ネットワーク結合が正常である限り、ローカルストレージシステムにより同期化される。

40

図 4 A - 4 E は、ローカルストレージシステムが障害になった場合に、各ストレージシステムで保有している受信されたデータメッセージに関するヒストリ情報を用いて、本発明によるリモートストレージシステムのデータイメージの同期方法を一般的に説明する。

図 4 A は、ローカルストレージシステム 4 0 4 が、リモートコピーキュー 4 0 4 a に 3 つのデータメッセージ (1 , 2 , 3) を持っている正常時の状態を示す。そのデータメッセージ中の二つ (1 , 2) がリモートストレージシステム 4 0 5 に送信され受信されているが、3 番目のデータメッセージ 3 は、当該リモートストレージシステム 4 0 5 に送信中である。しかしながら、リモートストレージシステム 4 0 6 には、データメッセージ 1 , 2 , 3 が一つも受け取られていなく、何れも送信準備中か、送信中である。

50

この状態で、即ち、データメッセージ 3 はリモートストレージシステム 405 で受信が完了しておらず、データメッセージ 1, 2, 3 のいずれもリモートストレージシステム 406 で受信が完了していない状態で、ローカルストレージシステム 404 が、障害になった場合の結果を図 4B に示す。この状態では、リモートストレージシステム 405, 406 間でデータメッセージは、リモートストレージシステム 405 のロールバックキュー 421 に蓄えられている、最新のデータメッセージ 1, 2 だけが相違している。

この状態での、本発明によれば、二つのストレージシステム 405, 406 により保持されているデータイメージを同期させる方式の一つの技術は、リモートストレージシステム 406 にも、データメッセージ 1, 2 を受け取らせ、リモートストレージシステム 406 が保有しているデータイメージを更新させることを必要とする。

10

かくして、図 4C に示すように、データイメージの相違はリモートストレージシステム 405 に保持される二つのデータメッセージ 1, 2 にあることを検出した上で、これらのデータメッセージをロールフォワード同期のためにリモートストレージシステム 406 に転送する。反対に、ロールバック方式の同期では、リモートストレージシステム 405 に、保有しているデータメッセージ 1, 2 を破棄させる。このロールバック同期方式を図 4E に示し、以下に論ずる。

図 4D は、システム中任意の場所に存在しえるソフトウェアマネージャ 407 により実行されるロールフォワード同期を示す。このようにこのマネージャはシステムの他の部分とは分離していながら、ローカルストレージシステムや、リモートストレージシステム 405, 406 にアクセス可能であってもよいし、全体のストレージシステムの一部を形成してもよい。

20

実装方法にかかわらず、マネージャ 407 は、プライマリストレージシステム 404 の障害検出を契機に始まる、リモートストレージシステム 405, 406 間の通信に立ち入る。この通信の結果により、マネージャ 407 は、プライマリストレージシステム 404 が倒れた時点の、更新情報の受信の差異をリモートストレージシステム間について調査する。図 4D の例では、マネージャ 407 は、リモートストレージシステム 405 が、リモートストレージシステム 406 が持っていないデータメッセージ 1, 2 を受信しており、データメッセージ 1, 2 はストレージシステム 405 のロールバックキューの中にあると知らされる。

そこで、このマネージャ 407 はリモートストレージシステム 405 にアクセスし、メッセージ 1, 2 に関するデータをリモートストレージシステム 406 に転送し、リモートストレージシステム 406 のデータイメージをリモートストレージシステム 405 のデータイメージと同期させる。

30

代わりに、マネージャ 407 は、リモートストレージシステム 405 のロールバックキューのデータメッセージを破棄させることも出来る。この場合、当該データメッセージは、中間キューには入っていないため、メッセージに対応するデータは書き込まれず、データイメージは変更されない。メディアこの後者の同期方式はロールバック同期と呼ばれ、リモートストレージシステム 405 のデータイメージをリモートストレージシステム 406 のデータイメージに同期させるものである。

図 4E は、ロールバック同期処理の他のやり方を説明する。前と同様に、プライマリストレージシステム 404 が、リモートストレージシステムのデータイメージを更新するために転送すべき、データメッセージ 1, 2, 3 を保有したまま、障害になったとする。このとき、リモートストレージシステム 405 のみがデータが更新されて、データメッセージ 1, 2 のみ保有しており、リモートストレージシステム 406 は全く更新されず、データメッセージ 1, 2, 3 の何れも保有していないとする。

40

従って、プライマリ（ローカル）ストレージシステム 404 がこの状態で障害になった場合は、両リモートストレージシステム 405, 406 のデータイメージは、データメッセージ 1, 2 だけが相違している。本発明のロールバック処理では、それぞれの保有するデータイメージを同期させるために、データメッセージ 1, 2 のデータをリモートストレージシステム 405 からリモートストレージシステム 406 にフォワード転送するのではな

50

く、リモートストレージシステム 405 が保有しているデータメッセージを破棄させる。図 5 は、ローカルストレージシステム 104 (図 1) が、ホスト 101 (図 1) から受信した更新データを各リモートストレージシステムに転送する、リモートコピー処理のために行う主要ステップを説明する。既に説明したように、ローカルストレージシステム 104 の保有するデータイメージを更新させるホストプロセッサ 101 からの書き込み I/O 要求は、リモートストレージシステムの保有するデータイメージのミラーに就いても同様な更新を必要とする。

この書き込み I/O 要求の発生により、アサインされた順序番号及び、この要求データへのポインタを含む、対応するコマンドエントリが形成される。このコマンドエントリは、全てのデータが受信され、完了報告がホストプロセッサ 101 に返されるまで、ワーキングキュー 110 に収容されている。その後、このコマンドエントリはロールバックキュー 111 に移される。その後、このロールバックキュー 111 が満杯になるか、フラッシュコマンド (後述する) が受信されると、このコマンドエントリは中間キュー 112 に移される。中間キュー 112 に存在している間に、この要求に対応するデータがストレージメディア 142 に書き込まれる。

CPU 133 上で走行している、リモートコピープロセスは、周期的にリモートコピーキューの内容をチェックして、リモートストレージシステム 105, 106 の各々に更新データをコピーするために受信した I/O 要求に対してデータメッセージの形成が必要なコマンドエントリを見つける。図 5 のステップ 501 にて、ローカルストレージシステム 104 は各リモートストレージシステムがデータメッセージの受付が出来るか否かをチェックする。

このチェックは、RC 状態テーブル 114 を参照して行う。このテーブルでは、各リモートストレージシステムが、どんなメッセージを受信して完了報告したか、完了報告していないか、及びリモートストレージシステムの各キューの受信可能サイズをローカルストレージシステムに伝える。このテーブルによりローカルストレージシステムは、特定のストレージシステムが別のデータメッセージや関連データの受信するスペースがあるか否かを判断する。もし否であったら、ステップ 501 は次の処理に移る。リモートストレージシステム自身が、これ以上データメッセージを受け取れないことの表示、例えば、SCSI 等で定義された "BUSY (0x08) : 使用中" や "キュー満杯 (0x28)" 等の状態信号を返答することも出来る。

もし、対象リモートストレージシステムがデータメッセージを受信するのに十分なスペースを持っていることが分かれば、ローカルストレージシステムは、リモートストレージシステム (例えばリモートストレージシステム 105) にデータメッセージの形式で RC (リモートコピー) コマンドを発行し、これに続いてステップ 503 で関連するデータを送信する。

ローカルストレージシステム 104 は、その後、データメッセージを受信したか否かの状態報告を待ち (ステップ 504)、ローカルストレージシステムが報告を受信すると、ステップ 505 で、全てのリモートストレージシステムがデータメッセージの更新が済んだかのチェックを行う。未了のリモートストレージシステムが存在すれば、処理はステップ 506 a に移り、対象となるリモートストレージシステムがデータイメージを受信したことを示すように RC 状態テーブル 114 を更新して、次のリモートストレージシステムがデータメッセージを受け取るために、ステップ 501 に戻る。

しかしながら、ステップ 505 にて全てのリモートストレージシステムがデータメッセージを受信したことが判明すれば、当該データメッセージ (コマンドエントリ) は、ステップ 506 で RC キューから削除され、ステップ 507 で RC 状態テーブル 114 が更新され、このデータメッセージのための処理は終了する。

図 6 は、リモートストレージシステムがリモートコピープロセスのデータメッセージを受信するステップを一般的に示す。ステップ 601 でリモートストレージシステムはデータメッセージを受信し、ステップ 602 で当該データメッセージを受け取るための十分なキュー資源がアサイン可能か、即ち受信のスペースがあるかを判定する。

10

20

30

40

50

更に、このステップ602の判定は、他のリモートストレージシステムのキュー内容に比較して、自らのキュー（即ち、ロールバック、中間、ライトヒストリキュー）の状態の分布、即ちアラインメント（後述される）に基づいて実施される。もし、ステップ602で、この時点では、当該データの受け取りは不可能と判定されたら、ステップ602からステップ606に移り、このリモートストレージシステムは、ローカルストレージシステム104に対して、“使用中”の応答を返し、この受信処理を終える。ローカルストレージシステムは、この応答により後の再試行の必要性を確認する。

もし、ステップ602で、受付可能と判定されたら、ステップ603でデータを受信し、ステップ604でリモートストレージシステムはデータ転送の状態を返す。続いて、ステップ605でこのデータメッセージは、メッセージ及びデータ受信用のワーキングキューから、ロールバックキューに移される。

10

図7は、ローカルストレージシステムで（又は、図3Eのデ이지チェーン接続での中間サイトのリモートストレージシステムで）障害が検出された場合の同期プロセスの処理ステップを示す。ステップ701で障害が検出される。障害検出は数々の方法でなされる。例えば、ストレージシステムは、互いに、ハートビート信号を他のストレージシステムに送信し、これが来なければ当該システムは障害とみなす；或いはストレージシステムは自己診断可能で、自システムの障害を他システムに通報通信することが出来る；或いはその他の任意の慣用的障害検知技術を用いることが出来る。

もし、障害が検出されると、他の生存システム、例えば、図2のリモートストレージシステム105、106は互いに交信し、誰が新しいリモートコピーマネージャになれるかを判定する。選択されたストレージシステムが、残りのすべての生存システムのデータメッセージを同一状態に同期するためのリモートコピー処理をコントロールする。あるいは、リモートコピーマネージャは、システム管理者により事前定義されている場合もある。

20

いずれにせよ、選択されると、リモートコピーマネージャは、ローカルストレージシステムとなり、ステップ703で生存ストレージシステムで維持されている、各種キューの構造と内容に関する情報を集める。この情報には、ロールバックキューとライトヒストリキューの範囲（エントリ数）と、どれだけの数のエントリがデータメッセージを有するかの情報が含まれる。例えば、図2について、ローカルストレージシステム104が障害になり（ステップ701）、リモートストレージシステム105がステップ702でリモートコピーマネージャとして選択されると、リモートストレージシステム106は、自分のロールバックキュー131にはデータメッセージ8、9、10、11が、ライトヒストリキュー133にはデータメッセージ3、4、5、6が保有されていることを報告する。

30

リモートストレージシステム106の中間キューの内容は、障害が検出されるとリモートストレージシステムは優先的に中間キューを空にして、それにより該当データがストレージメディアに書き込まれるようにするので、短時間のうちにライトヒストリキューに付け加えられる。

ストレージシステムの一つに障害が起きた後、生存ストレージシステムのリモートコピー環境に関する情報が得られた場合、選択されたリモートコピーマネージャは、ロールバックか、ロールフォワードかの何れの同期方式を選択するかを決める。これは、典型的には、リモートコピー環境の管理者又はユーザによって、障害の場合にストレージシステムを同期させる最良の方式に関するフラグをセットしておく形で、事前決定されている。

40

もし、ステップ704で、ロールバック処理が選択されたら、ステップ705でリモートコピーマネージャはすべてのストレージシステムが共有する最も大きな順序番号を持つデータメッセージを決定する。図2の例では、順序番号9のデータメッセージがこれに該当する。続いて、ステップ706にて、選択されたリモートコピーマネージャは、ロールバックコマンドを他の全てのストレージシステムに対して発行し、受信後順序番号9を超えるデータメッセージを破棄させる。かくして、図2に戻って、リモートストレージシステム106はこのロールバックコマンド受領により、順序番号10、11のデータメッセージを破棄する。この処理はステップ715で終了する。

一方において、ステップ704にて、ロールフォワードが選択された場合には、ステップ

50

707でロールフォワードの順序番号を決定する。この処理では、リモートコピー処理が、各種のロールバックキューとライトヒストリキューの内容を比較して、最新のデータメッセージを保有するストレージシステム決定する。かくして図2に於いて、選択されたりリモートコピーマネージャが、リモートストレージシステム105の場合は、リモートストレージシステム106が順序番号10, 11のメッセージを持っており、自分は持っていないことを認識する。

そこで、各システムのデータイメージの同期を計るために、リモートストレージシステム106の順序番号10, 11のデータメッセージを自システムに転送するリモートコピー処理を起動する。かくして、ステップ708にて、選択されたりリモートコピーマネージャは、それが最新のストレージシステムだと判定すると、ステップ709にて、最新データを有するストレージシステムから必要データを取得する。このことは、選択されたRCマネージャ自身にも起こり得るし、他のストレージシステムの一つにも起こり得る。いずれにせよ、選択されたRCマネージャが更新データを受領すると、ステップ710に移り、更新データを必要とする他のストレージシステムに対して更新データ又はその一部を選択的に転送してすべてのリモートストレージシステムのデータイメージの同期を完了する。この処理はステップ715で完結する。

もし、逆に、ステップ708にて、同期処理は、最新データメッセージを保有するストレージシステムにより実施され、当該ストレージシステムは選択されたRCマネージャではない場合(あるいは、他のストレージシステムが予め、同期システムとして、事前指定されている場合)は、RCマネージャは、どのストレージシステムがデータ更新を必要としているかを含む更新情報を、最新更新データ保有ストレージシステムに送信する。次にステップ712にて、更新データはデータメッセージの形で、このデータイメージの同期を必要とするすべてのストレージシステムに送信され、処理はステップ715で完了する。もし、(図4Eで説明され、既に議論されたように)ロールバック処理が採用された場合には、リモートストレージシステム105, 106(図1)のRB(ロールバック)キューは、アライン(調整)されている必要がある。アラインメント処理は、ローカルストレージシステム104で、CPU133上で走行しているリモートコピープロセスと(リモートストレージシステム105, 106で維持されている各種のキューの状態に関する情報を提供する)リモートコピー状態テーブル114を用いて、実行される。

例えば、あるリモートストレージシステムはローカルストレージシステムからのリモートコピーデータメッセージを受信できず、他のリモートストレージシステムは受信できるという事態が存在する。このような場合には同期目的のために、注意しなければならないことは、同期が必要なときに、同期が必ず成立するためには、一つのリモートストレージシステムのキューの中にあるデータメッセージと他のリモートストレージシステムのキューの中にあるデータメッセージとを比較したときに、少なくとも一つのデータメッセージが共通に保有(オーバーラップ)されている必要があることである。

図8, 9はすべてのキュー(特にRBキューと、ライトヒストリキュー)に就いての、アラインメント処理を説明する。図8では、リモートストレージシステム805/806のRBキュー821/831は2エントリの深さで、ライトヒストリキュー823/833はそれぞれ5エントリの深さの状態でのオーバーラップを説明する。(ワーキングキュー及び中間キューは、不必要に話を複雑化させるために図示されていない)。図8に於いて、リモートストレージシステム805は、ライトヒストリキュー823にデータメッセージ6, 7, 8, 9, 10が、RBキュー821にデータメッセージ11, 12が登録されている。

リモートストレージシステム805は、更にメッセージ13の受信途上であり、ローカルストレージシステム(図8に図示されていない)は、更にデータメッセージ14を送る準備をしている。リモートストレージシステムがデータメッセージ13を受信し受け付けると、このメッセージ13は、RBキュー821に入り込む。これにより、データメッセージ11が、スペースを作るためにライトヒストリキューに移動して、データメッセージエントリ6がライトヒストリキューから排出されることになる。リモートストレージシステム

10

20

30

40

50

805は更に、順序番号14のデータメッセージを受け取れる状態になる。

同時にリモートストレージシステム806は、ライトヒストリキュー833にデータメッセージ1-5を持ち、RBキュー831にデータメッセージ6,7を持っており、データメッセージ8,9を受信中である。

この状態で、更に、リモートストレージシステム805がメッセージ14受信し受け付けたとすると、ライトヒストリキュー823から末尾にあるメッセージ7が脱落することになる。この状態で、リモートストレージシステム806がデータメッセージ8,9を受信し受け付ける前に、例えばローカルストレージシステムが障害になったとすると、データメッセージ7に関する情報がリモートストレージシステム805から欠落しているため、ロールバック方式を用いて二つのリモートストレージシステムを同期させることができない結果になってしまう。

10

このことより、ローカルストレージシステム(図8には図示されていない)は、データメッセージの受信状態及びリモートストレージシステム805,806のキューの保有状態に関する、リモートコピー状態テーブル114(図2)からの情報により、リモートストレージシステム806が少なくともデータメッセージ8を受信し受け付けられる迄は、データメッセージ14の送信を抑制することによって両者のアライメントを維持する必要がある。

図9は、ローカルストレージシステム104で採用され、定期的に発行されるフラッシュコマンドを用いたアライメント処理を説明する。図9では、ローカルストレージシステム104は、データメッセージ6を指定して、フラッシュコマンドを発行する。これにより、両リモートストレージシステム805,806とも、データメッセージ6に後続するすべてのデータメッセージ(7,8,...)の保有を確実にさせる。しかしながら、リモートストレージシステム805は依然として、データメッセージ14は受信不可能で“使用中”状態を返答する。

20

【発明の効果】

本発明により、従来の技術に勝る数々の便益が実現される。

【図面の簡単な説明】

【図1】図1は、本発明を実装する、ローカル及び複数のリモートストレージシステムを有するデータ処理システムのブロックダイアグラムである。

【図2】図2は、ローカルストレージからリモートストレージシステムに送られた更新データのヒストリ情報を保持するために、ローカル及び複数のリモートストレージシステムの各々で実装されるキュー構造を説明するダイアグラムである。

30

【図3A】図3Aは、ローカル及びリモートストレージシステム間の結合方式に対応して、発生し得る異なった障害モードを、図で説明する。

【図3B】図3Bは、ローカル及びリモートストレージシステム間の結合方式に対応して、発生し得る異なった障害モードを、図で説明する。

【図3C】図3Cは、ローカル及びリモートストレージシステム間の結合方式に対応して、発生し得る異なった障害モードを、図で説明する。

【図3D】図3Dは、ローカル及びリモートストレージシステム間の結合方式に対応して、発生し得る異なった障害モードを、図で説明する。

40

【図3E】図3Eは、ローカル及びリモートストレージシステム間の結合方式に対応して、発生し得る異なった障害モードを、図で説明する。

【図4A】図4Aは、ストレージシステムの一つ(例えばローカルストレージシステム)が障害になり、他の生存ストレージシステムがデータイメージを同期させるための、異なったシナリオを図で説明する。

【図4B】図4Bは、ストレージシステムの一つ(例えばローカルストレージシステム)が障害になり、他の生存ストレージシステムがデータイメージを同期させるための、異なったシナリオを図で説明する。

【図4C】図4Cは、ストレージシステムの一つ(例えばローカルストレージシステム)が障害になり、他の生存ストレージシステムがデータイメージを同期させるための、異な

50

ったシナリオを図で説明する。

【図４Ｄ】図４Ｄは、ストレージシステムの一つ（例えばローカルストレージシステム）が障害になり、他の生存ストレージシステムがデータイメージを同期させるための、異なったシナリオを図で説明する。

【図４Ｅ】図４Ｅは、ストレージシステムの一つ（例えばローカルストレージシステム）が障害になり、他の生存ストレージシステムがデータイメージを同期させるための、異なったシナリオを図で説明する。

【図５】図５は、ローカルストレージのデータイメージになされたデータ変更をリモートストレージシステムにデータメッセージを送信することによってコピーする、リモートコピー処理の主要ステップを説明する。

10

【図６】図６は、ローカルストレージシステムからのデータメッセージの受領に応じてリモートストレージシステムが行う主要ステップを説明する。

【図７】図７は、ローカルストレージシステムでの障害が検出された時に、リモートストレージシステムが行う主要ステップを説明する。

【図８】図８は、各リモートストレージシステム間での情報キューのオーバーラップを維持するために、ローカルストレージシステムで発行される、アラインメント処理及びフラッシュ処理を図で説明する。

【図９】図９は、各リモートストレージシステム間での情報キューのオーバーラップを維持するために、ローカルストレージシステムで発行される、アラインメント処理及びフラッシュ処理を図で説明する。

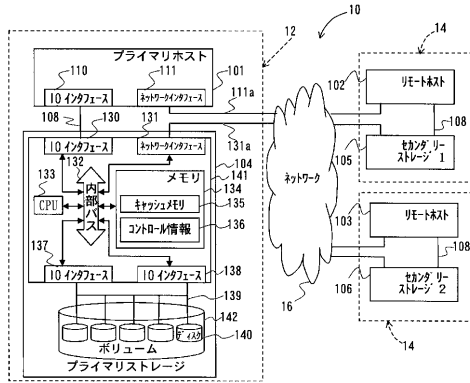
20

【符号の説明】

１６・・・ネットワーク、１０１・・・プライマリホスト、１０２・１０３・・・リモートホスト、１０４・・・プライマリストレージ、１０５・１０６・・・セカンダリストレージ、１１０・１３０・１３７・１３８・・・Ｉ／Ｏインタフェース、１１１・１３１・・・ネットワークインタフェース、１３２・・・内部バス、１３３・・・ＣＰＵ、１３４・・・メモリ、１３５・・・キャッシュメモリ、１３６・・・コントロール情報、１４０・・・ディスク、１４２・・・ボリューム

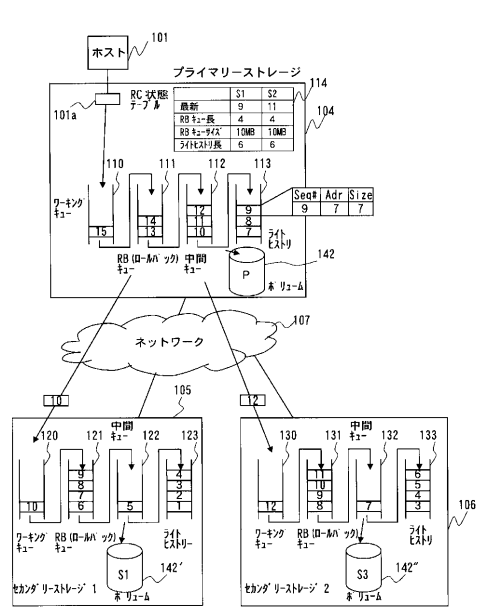
【 図 1 】

【図1】ローカル及び複数のリモートストレージシステムを有するデータ処理システム



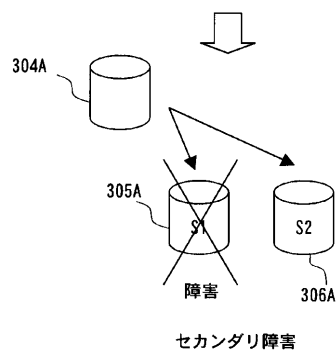
【 図 2 】

【図2】 ローカル及び複数のリモートストレージシステムでのキュー構造



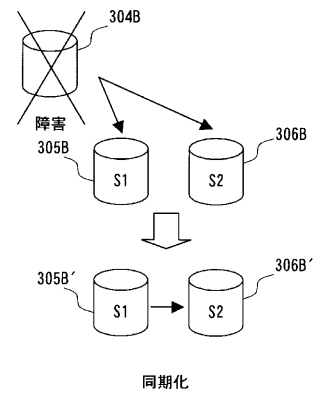
【 図 3 A 】

【図 3 A】



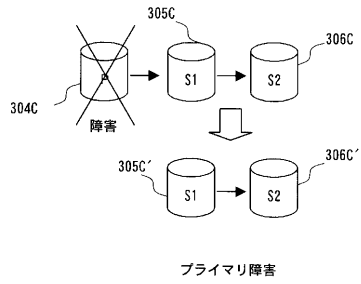
【 図 3 B 】

【図 3 B】



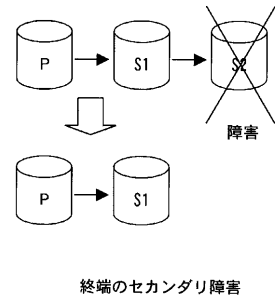
【図 3 C】

【図 3 C】



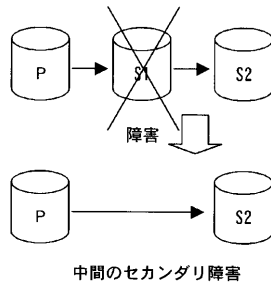
【図 3 D】

【図 3 D】



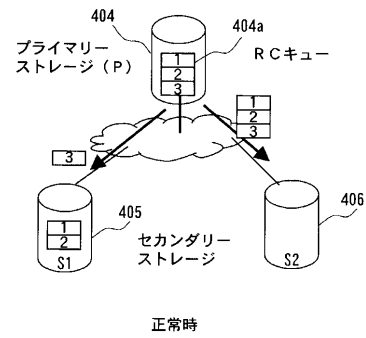
【図 3 E】

【図 3 E】



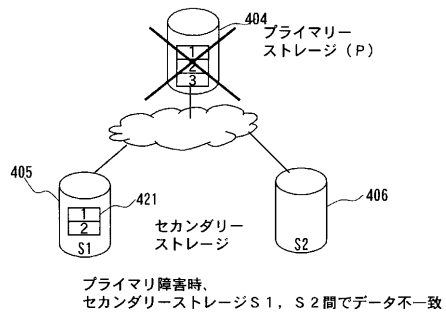
【図 4 A】

【図 4 A】



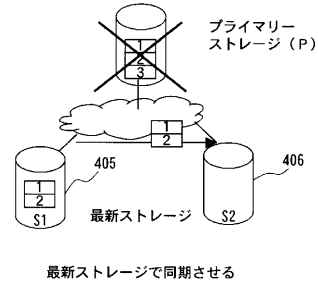
【図 4 B】

【図 4 B】



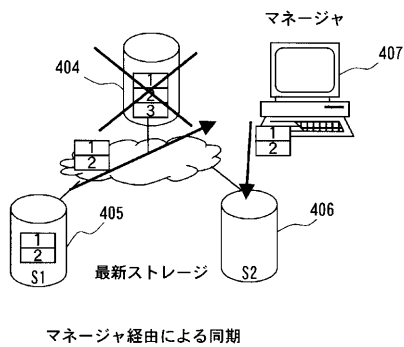
【図 4 C】

【図 4 C】



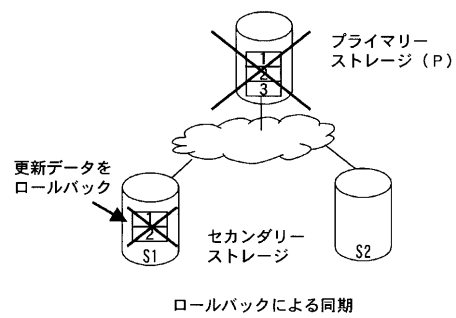
【図 4 D】

【図 4 D】



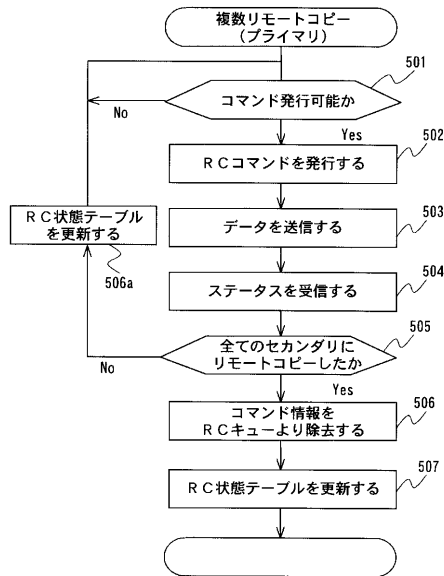
【図 4 E】

【図 4 E】



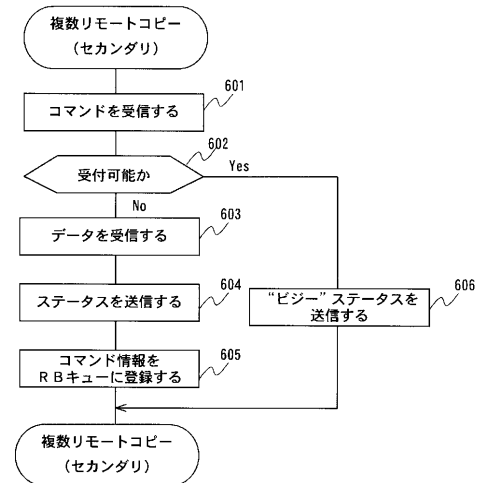
【図 5】

【図 5】



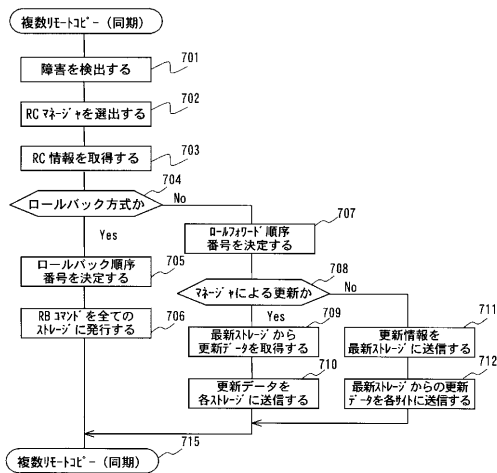
【図 6】

【図 6】



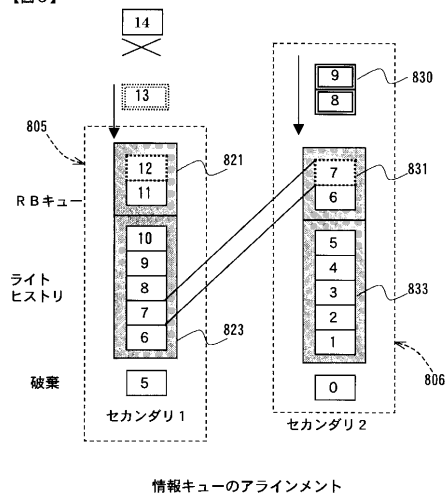
【図 7】

【図 7】

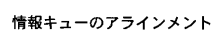


【図 8】

【図 8】



【图9】



フロントページの続き

(51)Int.Cl.

F I

G 0 6 F 13/10 3 4 0 A

G 0 6 F 15/00 3 1 0 B

G 0 6 F 15/00 3 2 0 D

(56)参考文献 特開 2 0 0 0 - 1 3 2 4 4 3 (J P , A)

特開 2 0 0 0 - 0 2 0 3 6 6 (J P , A)

特開 2 0 0 3 - 1 2 2 5 0 9 (J P , A)

特開平 1 1 - 2 9 6 4 2 2 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

G06F 3/06- 3/08

G06F 12/00

G06F 13/00-13/18