

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4549709号
(P4549709)

(45) 発行日 平成22年9月22日(2010.9.22)

(24) 登録日 平成22年7月16日(2010.7.16)

(51) Int.Cl.

F I

G 0 6 F 3/06 (2006.01)

G 0 6 F 3/06 3 0 4 F

請求項の数 13 (全 22 頁)

(21) 出願番号 特願2004-85032(P2004-85032)
 (22) 出願日 平成16年3月23日(2004.3.23)
 (65) 公開番号 特開2005-275582(P2005-275582A)
 (43) 公開日 平成17年10月6日(2005.10.6)
 審査請求日 平成19年2月21日(2007.2.21)

(73) 特許権者 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74) 代理人 100080001
 弁理士 筒井 大和
 (72) 発明者 武藤 義章
 神奈川県小田原市中里322番2号 株式
 会社日立製作所 R A I Dシステム事業部
 内
 (72) 発明者 竹内 久治
 神奈川県小田原市中里322番2号 株式
 会社日立製作所 R A I Dシステム事業部
 内

審査官 木村 雅也

最終頁に続く

(54) 【発明の名称】 ストレージ装置

(57) 【特許請求の範囲】

【請求項 1】

それぞれがキャッシュメモリを含み、それぞれが少なくとも1つのディスク装置と組み合わされ、少なくとも1つのホストコンピュータと接続された複数の正ディスク制御装置と、キャッシュメモリを含み、複数の前記正ディスク制御装置からリモートコピー処理によって送信されたデータを受信するよう構成された副ディスク制御装置と、前記副ディスク制御装置に接続されたディスク装置とを備え、

前記正ディスク制御装置のそれぞれのキャッシュメモリは、対応する前記ディスク装置へ書き込まれる更新データを格納する第1の領域と前記副ディスク制御装置に転送すべき更新データを格納する第2の領域を有し、

それぞれの前記正ディスク制御装置は、前記ホストコンピュータからの更新データを受信して、前記副ディスク制御装置に転送すべき更新データとして、前記キャッシュメモリの前記第2の領域に格納し、その格納された更新データを前記副ディスク制御装置に転送して、前記副ディスク制御装置へのリモートコピーを行うストレージ装置であって、

それぞれの前記正ディスク制御装置では、前記キャッシュメモリの前記第2の領域に格納された更新データに対して、予め設定された第1の基点と第2の基点間の規定時間の少なくとも1回の時間周期の間に上書きし、前記ホストコンピュータから送られた前記更新データに対し、少なくとも1回の時間周期の間に、前記更新データで定義された上書き対応レコードを記録し、上書きおよび格納された、更新データは共に前記副ディスク制御装置に転送され、

10

20

前記副ディスク制御装置では、複数の前記正ディスク制御装置から送られた更新データを時系列にデータ整合性を持つ一貫性の取れた実データとして扱い、

複数の前記正ディスク制御装置のそれぞれが、前記キャッシュメモリおよび制御メモリを含み、前記ホストコンピュータからのデータにタイムスタンプおよび受領番号を付与して前記キャッシュメモリの前記第2の領域に格納し、前記正ディスク制御装置のそれぞれの前記キャッシュメモリに格納された更新データに関して、レコードを上書きした際、その上書き回数の情報を前記制御メモリ内に格納し、

前記副ディスク制御装置では、前記タイムスタンプ、前記受領番号および前記上書き回数の情報に基づいて、全ての前記更新データを受領したかの確認を行うことを特徴とするストレージ装置。

10

【請求項2】

請求項1記載のストレージ装置において、

前記正ディスク制御装置のそれぞれのキャッシュメモリに格納された更新データに対して、同一レコードの更新データを上書きするモードおよび上書きしないモードを設け、前記ホストコンピュータからモード切替の指示があると、前記正ディスク制御装置のそれぞれのキャッシュメモリにモード切替の情報を格納して、前記モード切替の情報を前記副ディスク制御装置に転送し、

前記副ディスク制御装置では、前記正ディスク制御装置のそれぞれから転送された前記モード切替の情報に基づいて、動作モードを変更することを特徴とするストレージ装置。

【請求項3】

20

請求項2記載のストレージ装置において、

前記正ディスク制御装置のそれぞれの内部における前記更新データの処理状態に基づいて、前記モード切替の情報を、前記正ディスク制御装置のそれぞれのキャッシュメモリに格納することを特徴とするストレージ装置。

【請求項4】

請求項1記載のストレージ装置において、

前記タイムスタンプは、前記ホストコンピュータまたは複数の前記正ディスク制御装置のそれぞれにより付与され、前記受領番号は、複数の前記正ディスク制御装置のそれぞれにより前記ホストコンピュータからの更新データの受領情報に基づいて付与され、これらの情報は、前記制御メモリ内に格納されることを特徴とするストレージ装置。

30

【請求項5】

請求項1記載のストレージ装置において、

第2の基点は前記正ディスク制御装置で同期ポイントとして、前記正ディスク制御装置のキャッシュメモリに格納し、

前記副ディスク制御装置では、第1の基点と前記同期ポイント間の更新データを一貫性の取れた実データとして扱うことを特徴とするストレージ装置。

【請求項6】

請求項5記載のストレージ装置において、

前記同期ポイントは、前記ホストコンピュータからの同期要求により前記正ディスク制御装置のキャッシュメモリに格納されることを特徴とするストレージ装置。

40

【請求項7】

請求項1記載のストレージ装置において、

それぞれの前記正ディスク制御装置は、第1の基点と第2の基点間の更新データを一括して、前記副ディスク制御装置に転送することを特徴とするストレージ装置。

【請求項8】

請求項1記載のストレージ装置において、

前記正ディスク制御装置のそれぞれが、前記キャッシュメモリの前記第2の領域に格納された前記副ディスク制御装置に転送すべき更新データに対して、予め設定された第1の基点と第2の基点間の規定時間の間は同一レコードの更新データを上書きするよう構成されたことを特徴とするストレージ装置。

50

【請求項 9】

請求項 8 記載のストレージ装置において、

前記正ディスク制御装置と前記副ディスク制御装置との間で、前記第 1 の基点および前記第 2 の基点を同期させることを特徴とするストレージ装置。

【請求項 10】

請求項 1 記載のストレージ装置において、

前記正ディスク制御装置のそれぞれが、複数の時間周期の間に前記キャッシュメモリの前記第 2 の領域に格納されている上書きされた前記更新データに対して上書きし、それぞれの時間周期に前記ホストコンピュータから送られた前記更新データに対し、前記更新データによって前記時間周期に定義された上書き対応レコードを記録することで、それぞれの前記複数の時間周期に上書きされ格納された上書きデータが共に前記副ディスク制御装置に転送されることを特徴とするストレージ装置。

10

【請求項 11】

キャッシュメモリ、正ライトデータ受領部、正ライトデータ送信部およびディスクデータ送受信部を含み、少なくとも 1 つのホストコンピュータが接続され、そのそれぞれが自身のディスクデータ送受信部に接続された少なくとも 1 つのディスク装置と組み合わせられた複数の正ディスク制御装置と、

キャッシュメモリ、副ライトデータ受領部およびディスクデータ送受信部を含み、複数の正ディスク制御装置から送られたデータをリモートコピー処理によって受信するよう構成された副ディスク制御装置と、

20

前記副ディスク制御装置のディスクデータ送受信部に接続されたディスク装置とを備え、

前記正ディスク制御装置のそれぞれのキャッシュメモリは、前記ディスクデータ送受信部により前記ディスク装置へ書き込まれる更新データを格納する第 1 の領域と前記副ディスク制御装置に転送すべき更新データを格納する第 2 の領域を有し、

それぞれの前記正ディスク制御装置は、前記ホストコンピュータからの更新データを前記正ライトデータ受領部により受領し、前記ホストコンピュータから送られた前記更新データを受信して前記副ディスク制御装置に転送すべき更新データとして、前記キャッシュメモリの前記第 2 の領域に格納し、その格納された更新データを、前記正ライトデータ送信部により、前記副ディスク制御装置に転送して、前記副ディスク制御装置へのリモートコピーを行うストレージ装置であって、

30

前記それぞれの正ディスク制御装置では、前記キャッシュメモリの前記第 2 の領域に、予め設定された第 1 の基点と第 2 の基点間の規定時間の少なくとも 1 回の時間周期の間に更新データを上書きし、前記第 2 の基点を同期ポイントとして、前記正ディスク制御装置のキャッシュメモリに格納し、前記ホストコンピュータから送られた前記更新データに対し、少なくとも 1 回の時間周期の間に、前記更新データで定義された上書き対応レコードを記録することでその周期に上書きされ格納された更新データは、共に前記副ディスク制御装置に転送され、

前記副ディスク制御装置では、前記副ライトデータ受領部により前記正ライトデータ送信部から転送された更新データを受領し、その受領した更新データを前記副ディスク制御装置の前記キャッシュメモリに格納し、その格納された第 1 の基点と同期ポイント間の更新データを時系列にデータ整合性を持つ一貫性の取れた実データとして扱い、

40

複数の前記正ディスク制御装置のそれぞれが、キャッシュメモリおよび制御メモリを含み、前記ホストコンピュータからのデータにタイムスタンプおよび受領番号を付与して前記キャッシュメモリの前記第 2 の領域に格納し、前記正ディスク制御装置のそれぞれのキャッシュメモリに格納された更新データに関して、レコードを上書きした際、その上書き回数の情報を前記制御メモリ内に格納し、

前記副ディスク制御装置では、前記タイムスタンプ、前記受領番号および前記上書き回数の情報に基づいて、全ての前記更新データを受領したかの確認を行うことを特徴とするストレージ装置。

50

【請求項 1 2】

請求項 1 1 記載のストレージ装置において、

前記正ディスク制御装置のそれぞれが、前記キャッシュメモリの前記第 2 の領域に格納された前記副ディスク制御装置に転送すべき更新データに対して、予め設定された第 1 の基点と第 2 の基点間の規定時間の間は同一レコードの更新データを上書きするよう構成されたことを特徴とするストレージ装置。

【請求項 1 3】

請求項 1 1 記載のストレージ装置において、

前記正ディスク制御装置のそれぞれが、複数の時間周期にキャッシュメモリの前記第 2 の領域に前記上書きされた更新データを上書きし、それぞれの時間周期にホストコンピュータから送られた前記更新データに対し、前記時間周期に上書きされた前記更新データで定義された上書き対応レコードを上書きし、それによって複数の時間周期のそれぞれにおいて上書きされ格納された、前記上書きされた更新データが前記副ディスク制御装置に転送されることを特徴とするストレージ装置。

10

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、正副ディスク制御装置間で非同期のリモートコピーを行うストレージ装置に関し、特に、更新データの転送に適用して有効な技術に関するものである。

【背景技術】

20

【0002】

従来、ストレージ装置の正ディスク制御装置と副ディスク制御装置との間で非同期によるリモートコピーを行う際、正ディスク制御装置から副ディスク制御装置への更新データの転送としては、正ディスク制御装置へのライトデータを全て更新データとして転送していた。

【0003】

また、制御装置間の距離が拡大しても、性能の劣化を微小に抑えるために、正ディスク制御装置から、ライトデータの完了報告を返した後、直接副ディスク制御装置に送り、副ディスク制御装置は、受け取ったライトデータを不揮発メモリに格納することで、データ保証を行い、さらに、ある基準となる時刻を設け、この時刻以前の全てのライトデータを保証し、この時刻より後のライトデータは全て破棄できるようにした記憶制御装置があった（例えば、特許文献 1 参照）。

30

【特許文献 1】特開 2003 - 202962 号公報**【発明の開示】****【発明が解決しようとする課題】****【0004】**

しかしながら、従来のストレージ装置では、リモートコピーを行う際の正ディスク制御装置から副ディスク制御装置への更新データの転送は、正ディスク制御装置へのライトデータを全て更新データとして転送していたため、同一レコードに対して繰り返し上書きするアクセスパターンの場合には、転送データが増大し、I/O 遅延などが発生してしまうという問題があった。

40

【0005】

また、特許文献 1 記載のものも、同様に正ディスク制御装置へのライトデータを、副ディスク制御装置へ送り副ディスク制御装置の不揮発メモリに格納しているので、同一レコードに対して繰り返し上書きするアクセスパターンの場合には、転送データが増大し、I/O 遅延などが発生してしまうという問題があった。

【0006】

本発明の目的は、ストレージ装置において、リモートコピーを行う際の正ディスク制御装置から副ディスク制御装置への転送データのデータ量を低減させ、データの整合性も保証することができるストレージ装置を提供することにある。

50

【 0 0 0 7 】

また、データの整合性も保証する際の中断点の検出を副ディスク制御装置で容易に検出することのできるストレージ装置を提供することにある。

【課題を解決するための手段】

【 0 0 0 8 】

本発明によるストレージ装置は、キャッシュメモリおよび制御メモリを含む正ディスク制御装置と、正ディスク制御装置に接続されたディスク装置と、キャッシュメモリおよび制御メモリを含む副ディスク制御装置と、副ディスク制御装置に接続されたディスク装置とを備え、正ディスク制御装置のキャッシュメモリは、ディスク装置へ書き込まれる更新データを格納する第1の領域と副ディスク制御装置に転送すべき更新データを格納する第2の領域を有し、正ディスク制御装置に接続されたホストコンピュータからの更新データにタイムスタンプおよび受領番号を付与して、副ディスク制御装置に転送すべき更新データとして、キャッシュメモリの第2の領域に格納し、その格納された更新データを副ディスク制御装置に転送して、副ディスク制御装置へのリモートコピーを行うストレージ装置であって、正ディスク制御装置では、キャッシュメモリの第2の領域に格納された副ディスク制御装置に転送すべき更新データに対して、予め設定された第1の基点と第2の基点間の規定時間の間は同一レコードの更新データを上書きし、副ディスク制御装置では、第1の基点と第2の基点間の更新データを一貫性の取れた実データとして扱うものである。

10

【 0 0 0 9 】

また、本発明によるストレージ装置は、正ディスク制御装置のキャッシュメモリに格納された更新データに対して、同一レコードの更新データを上書きするモードおよび上書きしないモードを設け、ホストコンピュータからモード切替の指示があると、正ディスク制御装置のキャッシュメモリにモード切替の情報を格納して、モード切替の情報を前記副ディスク制御装置に転送し、副ディスク制御装置では、正ディスク制御装置から転送されたモード切替の情報に基づいて、動作モードを変更するものである。

20

【発明の効果】

【 0 0 1 0 】

本発明によれば、更新レコード単位の時系列にデータ整合性を取る非同期リモートコピーを行う際、同一レコードへの上書きI/Oにおいても、データの整合性を保証しつつ転送データ量の削減と処理オーバーヘッドの削減を図ることができる。

30

【 0 0 1 1 】

また、本発明によれば、正ディスク制御装置で同期ポイントを用いるので、副ディスク制御装置では同期ポイントを検出するだけで、中断点の検出を行うことができ、自由に同期ポイントを制御することができる。

【発明を実施するための最良の形態】

【 0 0 1 2 】

以下、本発明の実施の形態を図面に基づいて詳細に説明する。なお、実施の形態を説明するための全図において、同一の部材には原則として同一の符号を付し、その繰り返しの説明は省略する。

40

【 0 0 1 3 】

(実施の形態1)

<ストレージ装置の構成>

図1により、本発明の実施の形態1によるストレージ装置の構成について説明する。図1は本発明の実施の形態1によるストレージ装置の構成を示す構成図である。

【 0 0 1 4 】

図1において、ストレージ装置は、ホストコンピュータ120が接続された正ディスク制御装置(以下、正DKCという)100、正DKC100に接続されるディスク装置110、副ディスク制御装置(以下、副DKCという)200、副DKC200に接続されるディスク装置210から構成されている。

【 0 0 1 5 】

50

正DKC100は、正ライトデータ受領部101、正ライトデータ送信部102、キャッシュメモリ103、制御メモリ104、ディスクデータ送受信部108を含んでいる。

【0016】

キャッシュメモリ103には、正ライトデータ受領部101で受領された、ホストコンピュータ120からのライトデータ151が格納されている。キャッシュメモリ103に格納されたライトデータ151はその後ディスク装置110に書き込まれる。

【0017】

制御メモリ104には、ライトデータ151を管理するライトデータ管理情報105、ライトデータ151の受領番号(SEQ#)を管理するためのSEQ#マスタ106、正DKC100と副DKC200との間での基準点を管理するための基準点管理情報107が格納されている。

10

【0018】

ディスクデータ送受信部108は、ディスク装置110との間でデータの送受信を行う。

【0019】

正ライトデータ受領部101、正ライトデータ送信部102、キャッシュメモリ103、制御メモリ104およびディスクデータ送受信部108は、バスやスイッチなどにより接続されている。

【0020】

副DKC200は、副ライトデータ受領部201、キャッシュメモリ202、制御メモリ203、副ライトデータ一貫性保持部204、ディスクデータ送受信部208を含んでいる。

20

【0021】

キャッシュメモリ202には、正DKC100の正ライトデータ送信部102から送信され、副ライトデータ受領部201で受領された、ライトデータ151が格納されている。キャッシュメモリ202に格納されたライトデータ151はその後ディスク装置210に書き込まれる。

【0022】

制御メモリ203には、ライトデータ151を管理するライトデータ管理情報105、正DKC100と副DKC200との間での基準点を管理するための基準点管理情報107が格納されている。

30

【0023】

ディスクデータ送受信部208は、ディスク装置210との間でデータの送受信を行う。

【0024】

副ライトデータ受領部201、キャッシュメモリ202、制御メモリ203、副ライトデータ一貫性保持部204およびディスクデータ送受信部208は、バスやスイッチなどにより接続されている。

【0025】

本実施の形態では、このストレージ装置により、正DKC100と副DKC200との間で更新レコード単位のデータ整合性を取るリモートコピーを行っている。

40

【0026】

<ライトデータ管理情報>

次に、図2により、制御メモリ104、203内に格納されるライトデータ管理情報105について説明する。図2はライトデータ管理情報の一例を示す図である。

【0027】

図2において、ライトデータ管理情報105は、ライトデータ151に対応して作成され、ディスク装置アドレス130、ライトアドレス131、ライト時刻132、受領番号133、ライトデータポインタ134、転送必要ビット135から構成されている。

【0028】

50

ディスク装置アドレス 130 およびライトアドレス 131 は、ディスク装置 110, 210 上でのライトデータの位置を特定する情報であり、ホストコンピュータ 120 からライト要求 150 により指示される情報である。

【0029】

ライト時刻 132 の情報は、ホストコンピュータ 120 からのライト要求 150 内でライトデータ 151 に対してライト時刻のタイムスタンプ (T/S) 152 として付与された情報であり、ホストコンピュータ 120 がタイムスタンプ 152 を付与しない場合は、正 D K C 100 内のタイマを正 D K C 100 が付与する。

【0030】

受領番号 133 の情報は、正 D K C 100 が S E Q # マスタ 106 を参照して、ライトデータ 151 の受領順にライトデータ 151 に対して付与する情報である。

【0031】

ライトデータポインタ 134 は、対応するライトデータ 151 へのポインタの情報である。

【0032】

これらの情報はライトデータ 151 と共に、キャッシュメモリ 103, 202 にも格納され、ディスク装置 110, 210 へのライトデータ 151 の書き込み、正 D K C 100 から副 D K C 200 へのライトデータ 151 の転送、副 D K C 200 内でのライトデータ 151 の並び替えやライトデータ 151 に抜けがないかのチェックなどに利用される。

【0033】

転送必要ビット 135 は、正 D K C 100 から副 D K C 200 へライトデータ 151 を転送する必要があるかを示す情報であり、例えば、転送必要ビット 135 が ' O N ' の場合は、そのライトデータ 151 は正 D K C 100 から副 D K C 200 への転送が必要であるとして正 D K C 100 から副 D K C 200 へ転送し、転送した後に、そのライトデータ 151 に対する転送必要ビット 135 を ' O F F ' にする。

【0034】

< 正 D K C から副 D K C へのライトデータの転送動作 >

次に、本発明の実施の形態 1 によるストレージ装置の正 D K C 100 から副 D K C 200 へのライトデータの転送動作について説明する。図 3 は一般的なライトデータの転送動作を説明するための説明図、図 4 は実施の形態 1 によるストレージ装置のライトデータの転送動作を説明するための説明図、図 5 は正 D K C 側のライトデータをキャッシュメモリに格納する動作を示すフローチャート、図 6 は正 D K C 側の非同期周期処理の動作を示すフローチャート、図 7 は副 D K C 側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

【0035】

まず、本実施の形態の正 D K C から副 D K C へのライトデータの転送動作を説明する前に、一般的なライトデータの転送動作を説明する。

【0036】

一般的なライトデータの転送動作としては、図 3 に示すように、正 D K C 100 側では、ホストコンピュータ 120 からの更新 I O (図 3 の (1), (2), (3), (4), (5), (6), (7)) に対しホストコンピュータ 120 が付与したタイムスタンプ 152 と正 D K C 100 が付与する受領番号 133 とを更新データのライトデータ 151 と一緒に、副 D K C 200 へ転送すべき更新データ (以下、S i d e F i l e という) としてディスク装置 110 に書き込むデータとは別にキャッシュメモリ 103 に格納し、副 D K C 200 へ転送する前に更新 I O を完結させる。その後、正 D K C 100 は副 D K C 200 へホスト I O とは非同期に S i d e F i l e を転送する。

【0037】

副 D K C 200 では受領した S i d e F i l e からライト時刻 132、受領番号 133 を元に時系列並び替えを行い、確定時刻 (中断点) (例えば、図 3 の副 D K C 側の B' まで) までのデータ一貫性を保証する。

10

20

30

40

50

【0038】

この動作では、正DKC100において同一レコードに対し繰り返し上書きするようなアクセスパターン（例えば、図3のA、A'、A''など）においても、全て更新IOとして、SideFileに格納し、副DKC200へ転送している。

【0039】

本実施の形態では、図4に示すように、基点1から基点2までの規定時間の間で、正DKC100のキャッシュメモリ103のSideFile上の副DKC200への未転送データのレコードに対する同一レコードの上書きを可能としている（例えば、図4のBレコードに対するB'の更新データ）。

【0040】

このように、規定時間の間、SideFile上のレコードに対する上書きを可能とすることで、転送データ量の削減と処理オーバーヘッドの削減が可能となる。

【0041】

正DKC100ではSideFileを順次副DKC200へ転送してもよいが、未転送のSideFile上のレコードに更新IOがあった場合、上書きを行う。副DKC200で受領したSideFileは規定時間内はキャッシュメモリ202上に保留しておき、規定時間内の全てのデータを受領したら、当該SideFileを実データとして扱う。

【0042】

なお、正DKC100と副DKC200での基点は、正DKC100と副DKC200との間でリモートペアを作成する際の最初の動作で決定される。

【0043】

この正DKC100側の動作としては、図5に示すように、まず、ホストコンピュータ120からの更新IOがあると、更新IOのタイムスタンプ152が現在の基点2（基点1+規定時間）よりも大きいか否かを判断する（S100）。そして、S100で基点2よりも大きいと判断されると規定時間内ではないので、更新IOの更新データをSideFileへ追加し（S103）、S100で基点2以下と判断されるとSideFileに更新IOのレコードが存在するか否かを判断する（S101）。

【0044】

そして、S101でSideFileに更新IOのレコードが存在しないと判断されると上書きするレコードがないので更新IOの更新データをSideFileへ追加し（S103）、S101でSideFileに更新IOのレコードが存在すると判断されるとそのSideFileのレコードのデータに対して更新IOの更新データを上書きする（S102）。

【0045】

また、正DKC100側では、図6に示すように、非同期周期処理として、SideFileの先頭から副DKC200へデータ転送処理を行っている（S110）。

【0046】

また、副DKC200側の動作としては、図7に示すように、正DKC100からSideFileを受領し、副DKC200側のSideFileとして格納した後、タイムスタンプ152が現在の基点2（基点1+規定時間）よりも大きいか否かを判断する（S120）。そして、S120で基点2よりも大きいと判断されるまで正DKC100からSideFileを受領し、S120で基点2よりも大きいと判断されると、基点1から基点2の間の更新データについてその受領番号133にもれがないかを判断する（S121）。

【0047】

そして、S121で受領番号133にもれがあった場合は基点1から基点2の間の全ての更新データを正DKC100から受領していないので、再度、正DKC100からSideFileを受領し、副DKC200側のSideFileとして格納する。

【0048】

10

20

30

40

50

そして、S 1 2 1で受領番号1 3 3にもれがないと判断されると、基点1から基点2のデータを一貫性の取れた実データとして、キャッシュメモリ2 0 2内で通常キャッシュに移す(S 1 2 2)。その後通常キャッシュのデータはディスク装置2 1 0に書き込まれ、一貫性の取れた実データとなる。

【0 0 4 9】

< 受領番号の確認動作 >

次に、図8により、図7のS 1 2 1での受領番号1 3 3の確認動作について説明する。図8は受領番号の確認動作を説明するための説明図である。

【0 0 5 0】

正DKC 1 0 0のキャッシュメモリ1 0 3のSide File上の同一レコードを更新データで上書きすると、副DKC 2 0 0側で全てのSide Fileを受領したかどうか判断が難しい。

【0 0 5 1】

そこで、本実施の形態では、Side Fileの管理情報として、タイムスタンプ1 5 2、受領番号1 3 3に加えて、レコードに対する上書き回数の情報を追加する。

【0 0 5 2】

例えば、図8に示す例では、ホストコンピュータ1 2 0からの更新IO(図8の(1)、(2)、(3)、(4)、(5))の内、図8の(4)に示す更新IO(B')はBレコードと同一レコードのデータなので、B'の更新データをBレコードに上書きする。この際、Bレコードの受領番号1 3 3はB'のものではなく、Bレコードの受領番号のままとし、Bレコードに対する上書き回数の情報として“1”を追加する。

【0 0 5 3】

副DKC 2 0 0側では、基点1から基点2までの規定時間内の到着レコード数(図8に示す例では3個)を、受領番号1 3 3で確認する際、上書き回数の情報(図8に示す例では1)を参照し、次の規定時間のSide Fileの最初の受領番号1 3 3(図8に示す例では5)と比較することにより、全てのSide Fileを受領したかどうかを確認することができる。すなわち、受領番号1 3 3による到着レコード数の情報と、各レコードの上書き回数の情報を加算した情報により、更新データが上書きされなかった場合の到着レコード数が確認できるので、その到着レコード数と次の規定時間のSide Fileの最初の受領番号1 3 3により、全てのSide Fileを受領したかどうかを確認することができる。

【0 0 5 4】

本実施の形態では、基点1から基点2までの規定時間の中で、正DKC 1 0 0のキャッシュメモリ1 0 3のSide File上の副DKC 2 0 0への未転送データの同一レコードに対する上書きをしているので、更新レコード単位の時系列にデータ整合性を取る非同期リモートコピーを行う際、同一レコードへの上書きIOにおいても、データの整合性を保証しつつ転送データ量の削減と処理オーバーヘッドの削減を図ることができる。

【0 0 5 5】

例えば、ホストコンピュータ1 2 0から4 KB / レコードの同一レコードに対し1, 000 iopsの更新がある場合、従来では4 KB / レコード × 1, 000 iops = 4 MBの転送が必要であったものが、本実施の形態では、最大4 KBに削減可能である。

【0 0 5 6】

(実施の形態2)

本実施の形態は、実施の形態1において正DKC 1 0 0と副DKC 2 0 0の両方で管理する基点1、基点2の代わりに、正DKC 1 0 0から上書きを中止したタイミングを示す同期ポイントを副DKC 2 0 0へ転送することにより、副DKC 2 0 0側では、基点の管理は行わないようにしたものである。

【0 0 5 7】

本実施の形態のストレージ装置の構成は実施の形態1と同じ構成である。

【0 0 5 8】

10

20

30

40

50

< 正 D K C から副 D K C へのライトデータの転送動作 >

次に、本発明の実施の形態 2 によるストレージ装置の正 D K C 1 0 0 から副 D K C 2 0 0 へのライトデータの転送動作について説明する。図 9 は実施の形態 2 によるストレージ装置のライトデータの転送動作を説明するための説明図、図 1 0 は正 D K C 側のライトデータをキャッシュメモリに格納する動作を示すフローチャート、図 1 1 は正 D K C 側の非同期周期処理の動作を示すフローチャート、図 1 2 はホストコンピュータからの同期ポイント指示の動作を示すフローチャート、図 1 3 は副 D K C 側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

【 0 0 5 9 】

本実施の形態では、図 9 に示すように、基点 1 から同期ポイント（図 9 の S のレコード）までの時間の間で、正 D K C 1 0 0 のキャッシュメモリ 1 0 3 の S i d e F i l e 上の副 D K C 2 0 0 への未転送データのレコードに対する同一レコードの上書きを可能としている（例えば、図 9 の B レコードに対する B ' の更新データ）。

10

【 0 0 6 0 】

このように、ある一定時間の間、S i d e F i l e 上のレコードに対する上書きを可能とすることで、転送データ量の削減と処理オーバーヘッドの削減が可能となる。

【 0 0 6 1 】

正 D K C 1 0 0 では S i d e F i l e を順次副 D K C 2 0 0 へ転送してもよいが、未転送の S i d e F i l e 上のレコードに更新 I O があった場合、上書きを行う。副 D K C 2 0 0 で受領した S i d e F i l e は、同期ポイントが転送されるまではキャッシュメモリ 2 0 2 上に保留しておき、同期ポイントが転送されるまでの全てのデータを受領したら、当該 S i d e F i l e を実データとして扱う。

20

【 0 0 6 2 】

なお、正 D K C 1 0 0 と副 D K C 2 0 0 での基点 1 は、正 D K C 1 0 0 と副 D K C 2 0 0 との間でリモートペアを作成する際の最初の動作で決定され、その後、正 D K C 1 0 0 では、実施の形態 1 と同様な規定時間により、基点 2 の管理を行っており、基点 2 付近での同期ポイントが副 D K C 2 0 0 へ転送されるようになっている。また、ホストコンピュータ 1 2 0 から任意に同期ポイントを指示することもでき、ユーザが自由に非同期コピーの同期を制御することも可能である。

【 0 0 6 3 】

30

この正 D K C 1 0 0 側の動作としては、図 1 0 に示すように、まず、ホストコンピュータ 1 2 0 からの更新 I O があると、更新 I O のタイムスタンプ 1 5 2 が現在の基点 2（基点 1 + 規定時間）よりも大きいかなかを判断する（S 1 3 0）。そして、S 1 3 0 で基点 2 よりも大きいと判断されると同期ポイントの転送時期として、S i d e F i l e に同期ポイントを指示し（S 1 3 3）、更新 I O の更新データを S i d e F i l e へ追加し（S 1 3 4）、S 1 3 0 で基点 2 以下と判断されると S i d e F i l e に更新 I O のレコードが存在するか否かを判断する（S 1 3 1）。

【 0 0 6 4 】

そして、S 1 3 1 で S i d e F i l e に更新 I O のレコードが存在しないと判断されると上書きするレコードがないので更新 I O の更新データを S i d e F i l e へ追加し（S 1 3 4）、S 1 3 1 で S i d e F i l e に更新 I O のレコードが存在すると判断されるとその S i d e F i l e のレコードのデータに対して更新 I O の更新データを上書きする（S 1 3 2）。

40

【 0 0 6 5 】

また、正 D K C 1 0 0 側では、図 1 1 に示すように、非同期周期処理として、S i d e F i l e の先頭から副 D K C 2 0 0 へデータ転送処理を行っている（S 1 4 0）。また、非同期周期処理として、現在時刻が現在の基点 2（基点 1 + 規定時間）よりも大きいかなかを判断し（S 1 4 1）、現在時刻が現在の基点 2（基点 1 + 規定時間）よりも大きい場合は、S i d e F i l e に同期ポイントを指示している（S 1 4 2）。

【 0 0 6 6 】

50

なお、図10のS133と、図11のS142での同期ポイントの指示は、正DKC100で管理している基点2の時間付近で動作するものであり、どちらかの指示により同期ポイントが指示され、同期ポイントが指示された後は、そのポイントが基点1として管理されるため2重に同期ポイントが指示されないようになっている。

【0067】

また、ホストコンピュータ120からの同期要求があった際の動作としては、図12に示すように、ホストコンピュータ120からの同期要求があると、その時点でSideFileに同期ポイントを指示する(S150)。

【0068】

また、副DKC200側の動作としては、図13に示すように、正DKC100からSideFileを受領し、副DKC200側のSideFileとして格納した後、同期ポイントのSideFileか否かを判断する(S160)。そして、S160で同期ポイントのSideFileと判断されるまで正DKC100からSideFileを受領し、S160で同期ポイントのSideFileと判断されると、基点1から同期ポイントの間の更新データの受領番号133を全て受領したか否かを判断する(S161)。

【0069】

そして、S161で受領番号133を全て受領していないと判断されると、再度正DKC100からSideFileを受領し、副DKC200側のSideFileとして格納する(S162)。

【0070】

そして、S161で受領番号133を全て受領したと判断されると、基点1から同期ポイントのデータを一貫性の取れた実データとして、キャッシュメモリ202内で通常キャッシュに移す(S163)。その後通常キャッシュのデータはディスク装置210に書き込まれ、一貫性の取れた実データとなる。

【0071】

本実施の形態では、基点1から同期ポイントまでの時間の間で、正DKC100のキャッシュメモリ103のSideFile上の副DKC200への未転送データの同一レコードに対する上書きをしているので、更新レコード単位の時系列にデータ整合性を取る非同期リモートコピーを行う際、同一レコードへの上書きIOにおいても、データの整合性を保証しつつ転送データ量の削減と処理オーバーヘッドの削減を図ることができる。

【0072】

例えば、ホストコンピュータ120から4KB/レコードの同一レコードに対し1,000iopsの更新がある場合、従来では4KB/レコード×1,000iops=4MBの転送が必要であったものが、本実施の形態では、最大4KBに削減可能である。

【0073】

さらに、副DKC200では基点を管理することなく同期ポイントを検出するだけで、中断点の検出を行うことができ、ユーザがホストコンピュータ120から自由に同期ポイントを制御することが可能となり、例えばDBMS(DataBase Management System)のコミット(ディスク書き込み)直後に同期ポイント指示を行えば、当該コミットが副DKC200へ確実に到着したことを確認できる。

【0074】

(実施の形態3)

本実施の形態は、実施の形態1において非同期での正DKC100から副DKC200への更新データの転送は行わずに、基点1～基点2までの規定時間内はSideFileを副DKC200へ転送せず、規定時間を超過した地点で当該SideFileを副DKC200へ一括転送するようにしたものである。

【0075】

本実施の形態のストレージ装置の構成は実施の形態1と同じ構成である。

【0076】

<正DKCから副DKCへのライトデータの転送動作>

10

20

30

40

50

次に、本発明の実施の形態 3 によるストレージ装置の正 D K C 1 0 0 から副 D K C 2 0 0 へのライトデータの転送動作について説明する。図 1 4 は実施の形態 3 によるストレージ装置のライトデータの転送動作を説明するための説明図、図 1 5 は正 D K C 側のライトデータをキャッシュメモリに格納する動作およびデータ転送動作を示すフローチャート、図 1 6 は副 D K C 側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

【 0 0 7 7 】

本実施の形態では、図 1 4 に示すように、基点 1 から基点 2 までの規定時間の間で、正 D K C 1 0 0 のキャッシュメモリ 1 0 3 の S i d e F i l e 上の副 D K C 2 0 0 への未転送データのレコードに対する同一レコードの上書きを可能としている（例えば、図 1 4 の B レコードに対する B ' の更新データ）。 10

【 0 0 7 8 】

このように、規定時間の間、S i d e F i l e 上のレコードに対する上書きを可能とすることで、転送データ量の削減と処理オーバーヘッドの削減が可能となる。

【 0 0 7 9 】

そして、正 D K C 1 0 0 では S i d e F i l e を基点 1 から基点 2 までの規定時間の間は副 D K C 2 0 0 へ転送せずに、S i d e F i l e 上のレコードに更新 I O があった場合、上書きを行う。副 D K C 2 0 0 で受領した S i d e F i l e は規定時間内はキャッシュメモリ 2 0 2 上に保留しておき、規定時間内の全てのデータを受領したら、当該 S i d e F i l e を実データとして扱う。 20

【 0 0 8 0 】

なお、正 D K C 1 0 0 と副 D K C 2 0 0 での基点は、正 D K C 1 0 0 と副 D K C 2 0 0 との間でリモートペアを作成する際の最初の動作で決定される。

【 0 0 8 1 】

この正 D K C 1 0 0 側の動作としては、図 1 5 に示すように、まず、ホストコンピュータ 1 2 0 からの更新 I O があると、更新 I O のタイムスタンプ 1 5 2 が現在の基点 2（基点 1 + 規定時間）よりも大きいか否かを判断し（S 1 7 0）、S 1 7 0 で基点 2 以下と判断されると S i d e F i l e に更新 I O のレコードが存在するか否かを判断する（S 1 7 1）。 30

【 0 0 8 2 】

そして、S 1 7 1 で S i d e F i l e に更新 I O のレコードが存在しないと判断されると上書きするレコードがないので更新 I O の更新データを S i d e F i l e へ追加し（S 1 7 3）、S 1 7 1 で S i d e F i l e に更新 I O のレコードが存在すると判断されるとその S i d e F i l e のレコードのデータに対して更新 I O の更新データを上書きする（S 1 7 2）。 30

【 0 0 8 3 】

また、S 1 7 0 で基点 2 よりも大きいと判断されると規定時間内ではないので、更新 I O の更新データを S i d e F i l e へ追加し（S 1 7 4）、基点 1 ~ 基点 2 の S i d e F i l e を一括して副 D K C 2 0 0 へデータ転送する（S 1 7 5）。 40

【 0 0 8 4 】

また、正 D K C 1 0 0 側では、図 1 6 に示すように、正 D K C 1 0 0 から基点 1 ~ 基点 2 の S i d e F i l e を受領し、副 D K C 2 0 0 側の S i d e F i l e として格納した後、基点 1 から基点 2 の間の更新データについてその受領番号 1 3 3 にもれがないかを判断する（S 1 8 0）。 40

【 0 0 8 5 】

そして、S 1 8 0 で受領番号 1 3 3 にもれがあった場合は基点 1 から基点 2 の間の全ての更新データを正 D K C 1 0 0 から受領していないので、再度、正 D K C 1 0 0 から S i d e F i l e を受領し、副 D K C 2 0 0 側の S i d e F i l e として格納する。

【 0 0 8 6 】

そして、S 1 8 0 で受領番号 1 3 3 にもれがないと判断されると、基点 1 から基点 2 の 50

データを一貫性の取れた実データとして、キャッシュメモリ 202 内で通常キャッシュに移す (S181)。その後通常キャッシュのデータはディスク装置 210 に書き込まれ、一貫性の取れた実データとなる。

【0087】

本実施の形態では、基点 1 から基点 2 までの規定時間の中で、正 DKC 100 のキャッシュメモリ 103 の Side File 上の副 DKC 200 への未転送データの同一レコードに対する上書きをしているので、更新レコード単位の時系列にデータ整合性を取る非同期リモートコピーを行う際、同一レコードへの上書き I/O においても、データの整合性を保証しつつ転送データ量の削減と処理オーバーヘッドの削減を図ることができる。

【0088】

例えば、ホストコンピュータ 120 から 4KB / レコードの同一レコードに対し 1,000 iops の更新がある場合、従来では 4KB / レコード \times 1,000 iops = 4MB の転送が必要であったものが、本実施の形態では、最大 4KB に削減可能である。

【0089】

さらに、Side File を正 DKC 100 から非同期で順次転送すると、同一レコードに更新 I/O があっても、当該データが既に転送済みである場合が想定され、規定時間内であっても再度更新データを転送しなければならないが、基点 1 ~ 基点 2 までの規定時間内の更新データを一括して転送することにより、Side File への上書き効率が向上し、転送データのさらなる削減が可能となる。

【0090】

また、副 DKC 200 では規定時間内データを一括受領することで、規定時間を意識することなく受領データを実データとして扱うことができる。

【0091】

(実施の形態 4)

本実施の形態は実施の形態 1 ~ 実施の形態 3 において、通信経路障害などで正 DKC 100 と副 DKC 200 との間の通信ができなくなり、リモートペアを一時停止した場合でもデータの整合性を保証できるようにしたものである。

【0092】

本実施の形態のストレージ装置の構成は実施の形態 1 と同じ構成である。

【0093】

< 通信経路障害時の動作 >

次に、本発明の実施の形態 4 によるストレージ装置の通信経路障害時の動作について説明する。

【0094】

通信経路障害等で正 DKC 100 と副 DKC 200 との間での通信ができなくなると当該リモートコピーペアを一時停止する必要がある。一時停止においても副 DKC 200 におけるデータ整合性は保証する必要があり、一時停止解除後は一時停止期間中の更新データも漏れなく副 DKC 200 に反映させなければならない。

【0095】

この場合、Side File 転送中かつ、規定時間内の Side File を全て副 DKC 200 に転送できなかった場合 (通信経路の障害等)、その期間内のデータは副 DKC 200 では破棄してよいが、正 DKC 100 では副 DKC 200 に送信してしまったデータを含めて未転送データとして扱う必要がある。

【0096】

そこで、本実施の形態では、図 2 に示すようなライトデータ管理情報 105 の転送必要ビット 135 を用いて、通信経路障害時の動作を制御する。

【0097】

まず、正 DKC 100 で更新データを Side File として格納した際、図 2 の転送必要ビット 135 を 'ON' とし、副 DKC 200 への転送が完了すると転送必要ビット 135 を 'OFF' とし、規定時間内の Side File データ全てを正常に副 DKC

10

20

30

40

50

200へ転送完了した場合、当該規定時間内のSideFileは正DKC100から消去可能とする。

【0098】

しかし、規定時間内のSideFileを全て送信できず、途中で転送に失敗した（通信経路障害等）場合、SideFile内の転送必要ビット135の‘ON’、‘OFF’の状態にかかわらず、SideFile内のデータ全てを副DKC200へ転送する必要のある差分データとして管理する。

【0099】

また、一時的な障害で即座に復旧できた場合は一時停止とはならず、転送必要ビット135の状態に従ってデータ転送を継続する。

10

【0100】

本実施の形態では、通常時は、ライトデータ管理情報105の転送必要ビット135を用いて、データ転送を行い、通信経路障害時には、転送必要ビット135の状態にかかわらずSideFile内のデータ全てを副DKC200へ転送する必要のある差分データとして管理するので、リモートコピーペアが停止しなければならない状態（通信経路障害等）となった場合、規定時間内のSideFileにおいて、副DKC200へ転送完了したデータも、差分データとして管理することが可能である。

【0101】

（実施の形態5）

本実施の形態は実施の形態1～実施の形態3において、SideFileの格納モードとして、SideFileを上書きする上書きモードと、SideFileを上書きしない従来モードを切り替えるようにしたものである。

20

【0102】

本実施の形態のストレージ装置の構成は実施の形態1と同じ構成である。

【0103】

<モード切替の動作>

次に、本発明の実施の形態5によるストレージ装置のモード切替の動作について説明する。

【0104】

図17は実施の形態5によるストレージ装置のモード切替の動作を説明するための説明図、図18は正DKC側のモード切替の動作を示すフローチャート、図19は副DKC側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

30

【0105】

本実施の形態では、図17に示すように、基点1から基点2までの規定時間の間で、ホストコンピュータ120からのモード切替指示の指定を可能としている（例えば、図17の（5）に示すホストコンピュータ120からのデータ）。

【0106】

このように、正DKC100側でモード切替指示の指定を可能とすることで、上書きモードと従来モードでの動作を自由に切り替えることが可能となる。

【0107】

そして、正DKC100ではモード切替を指示されたら、副DKC200へそれを通知し、その時刻以降にモード切替を行い、副DKC200では、正DKC100からのモード切替指示により、モード切替を行う。

40

【0108】

この正DKC100側の動作としては、図18に示すように、まず、ホストコンピュータ120からのIOがあると、IOがモード切替指示か否かを判断する（S190）。そして、S190でモード切替指示でないと判断されると、通常のIO処理を行う（S191）。

【0109】

また、S190でモード切替指示であると判断されると、モード切替指示をSideF

50

ileに格納し、SideFileを副DKC200へ転送することにより、モード切替指示を副DKC200へ通知し(S192)、モード切替を行う(S193)。

【0110】

また、副DKC200側の動作としては、図19に示すように、正DKC100からSideFileを受領し、副DKC200側のSideFileとして格納した後、タイムスタンプ152が現在の基点2(基点1+規定時間)よりも大きいか否かを判断する(S200)。そして、S200で基点2よりも大きいと判断されるまで正DKC100からSideFileを受領し、S200で基点2よりも大きいと判断されると、SideFile内にモード切替指示があるかを確認しモード切替か否かを判断する(S201)。

10

【0111】

そして、S201でモード切替と判断されると、当該DKCのモード切替を行い(S202)、S201でモード切替でないと判断されると、基点1から基点2の間の更新データについてその受領番号133にもれがないかを判断する(S203)。

【0112】

そして、S203で受領番号133にもれがあった場合は基点1から基点2の間の全ての更新データを正DKC100から受領していないので、再度、正DKC100からSideFileを受領し、副DKC200側のSideFileとして格納する。

【0113】

そして、S203で受領番号133にもれがないと判断されると、基点1から基点2のデータを一貫性の取れた実データとして、キャッシュメモリ202内で通常キャッシュに移す(S204)。その後通常キャッシュのデータはディスク装置210に書き込まれ、一貫性の取れた実データとなる。

20

【0114】

本実施の形態では、SideFileの格納モードとして、SideFileを上書きする上書きモードの場合は、副DKC200側で実データとして扱われるまで基点2や同期ポイントまで待つ必要があり、正DKC100とのタイムラグが長くなるが、正DKC100と副DKC200の一致性を優先したい場合には、SideFileを上書きしない従来モードに変更するように、SideFileの格納モードを動作中に変更することができるので、正DKC100と副DKC200のデータの一致性や正DKC100被災時の副DKC200でのデータ破棄量を調整することができる。

30

【0115】

なお、本実施の形態では、ホストコンピュータ120からのモード切替指示の例で説明したが、正DKC100の判断(例えば、更新IO過多によりSideFile満杯等)により上書きモード、従来モードの切替指示を行うようにしてもよい。

【0116】

(実施の形態6)

本実施の形態は、実施の形態5において、複数の正DKC100と単一の副DKC200とで構成しデータ整合性を保証する場合に、副DKCで正DKCごとの動作モードを管理するものである。

40

【0117】

本実施の形態のストレージ装置の構成は実施の形態1と同じ構成であり、複数の正DKC100と単一の副DKC200とが接続された構成となっている。

【0118】

<複数の正DKCからのモード切替>

次に、本発明の実施の形態6によるストレージ装置のモード切替の動作について説明する。

【0119】

図20は実施の形態6によるストレージ装置のモード切替の動作を説明するための説明図、図21は副DKC側のライトデータをキャッシュメモリに格納する動作を示すフロー

50

チャートである。

【0120】

本実施の形態では、図20に示すように、複数の正DKC100からのモード切替指示を副DKC200側で管理している。副DKC200側では、全ての正DKC100の動作モードが一致するまで、副DKC200側でSideFileの実データ化する動作モードを変えず、全ての正DKC100の動作モードが一致したとき、副DKC200側の動作モードを変更するようになっている。また、上書きモード、従来モード（同期モード）が混在するときは、副DKC200では全て上書きモードとしてSideFileを扱っている。

【0121】

この正DKC100側の動作としては、図18に示す実施の形態5の動作と同様に、まず、ホストコンピュータ120からのIOがあると、IOがモード切替指示か否かを判断する（S190）。そして、S190でモード切替指示でないと判断されると、通常のIO処理を行う（S191）。

【0122】

また、S190でモード切替指示であると判断されると、モード切替指示をSideFileに格納し、SideFileを副DKC200へ転送することにより、モード切替指示を副DKC200へ通知し（S192）、モード切替を行う（S193）。

【0123】

また、副DKC200側の動作としては、図21に示すように、正DKC100からSideFileを受領し、副DKC200側のSideFileとして格納した後、SideFile内にモード切替指示があるかを確認しモード切替か否かを判断する（S210）。

【0124】

そして、S210でモード切替と判断されると、当該DKCのモード切替を行い（S211）、S210でモード切替でないと判断されると、一時停止指示か否かを判断する（S212）。そして、S212で一時停止指示と判断されると正DKC100と副DKC200とのペアを一時停止する（S213）。

【0125】

また、S212で一時停止指示ではないと判断されると、現状は上書きモードか否かを判断し（S214）、S214で現状は上書きモードでないと判断されると従来モードである同期モードとして動作し、当該SideFile以前のデータ到着が確認できれば実データとして扱う（S215）。

【0126】

また、S214で現状は上書きモードであると判断されると、タイムスタンプ152が現在の基点2（基点1＋規定時間）よりも大きいと判断する（S216）。そして、S216で基点2よりも大きいと判断されるまで正DKC100からSideFileを受領し、S216で基点2よりも大きいと判断されると、基点1から基点2の間の更新データについてその受領番号133にもれがないかを判断する（S217）。

【0127】

そして、S217で受領番号133にもれがあった場合は基点1から基点2の間の全ての更新データを正DKC100から受領していないので、再度、正DKC100からSideFileを受領し、副DKC200側のSideFileとして格納する。

【0128】

そして、S217で受領番号133にもれがないと判断されると、当該正DKC100のSideFile完了フラグを‘ON’とし、他の正DKC100のSideFileが全て完了か否かを判断し（S219）、S219で他の正DKC100のSideFileが全て完了していないと判断されると、正DKC100からSideFileを受領し、副DKC200側のSideFileとして格納する。

【0129】

また、S 2 1 9で他の正D K C 1 0 0のS i d e F i l eが全て完了したと判断されると、全正D K C 1 0 0の指示モードは従来モードである同期モードか否かを判断し(S 2 2 0)、S 2 2 0で全正D K C 1 0 0の指示モードは同期モードであると判断されると、全D K C 1 0 0の実行モードを同期モードに変更し(S 2 2 1)、基点1から基点2のデータを一貫性の取れた実データとして扱う。また、S 2 2 0で全正D K C 1 0 0の指示モードは同期モードでないと判断されると、全正D K C 1 0 0の実行モードはそのまま変更せずに、基点1から基点2のデータを一貫性の取れた実データとして扱う。

【0 1 3 0】

本実施の形態では、複数の正D K C 1 0 0と単一の副D K C 2 0 0との間で動作モードを変更する際、副D K C 2 0 0側で動作モードを管理し、複数の正D K C 1 0 0の動作モードが一致したときに副D K C 2 0 0側の動作モードを変更するので、複数の正D K C 1 0 0間で新たな通信手段を用いなくても副D K C 2 0 0の制御で動作モードの不一致を許容することができる。また、動作モード変更を複数の正D K C 1 0 0の環境に容易に実装することができる。

10

【0 1 3 1】

なお、この実施の形態では、ホストコンピュータ1 2 0からのモード切替指示の例で説明したが、実施の形態5と同様に正D K C 1 0 0の判断(例えば、更新I O過多によりS i d e F i l e 満杯等)により上書きモード、従来モードの切替指示を行うようにしてもよい。

【図面の簡単な説明】

20

【0 1 3 2】

【図1】本発明の実施の形態1によるストレージ装置の構成を示す構成図である。

【図2】本発明の実施の形態1によるストレージ装置のライトデータ管理情報の一例を示す図である。

【図3】一般的なライトデータの転送動作を説明するための説明図である。

【図4】実施の形態1によるストレージ装置のライトデータの転送動作を説明するための説明図である。

【図5】本発明の実施の形態1によるストレージ装置の正D K C側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

【図6】本発明の実施の形態1によるストレージ装置の正D K C側の非同期周期処理の動作を示すフローチャートである。

30

【図7】本発明の実施の形態1によるストレージ装置の副D K C側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

【図8】本発明の実施の形態1によるストレージ装置の受領番号の確認動作を説明するための説明図である。

【図9】本発明の実施の形態2によるストレージ装置のライトデータの転送動作を説明するための説明図である。

【図10】本発明の実施の形態2によるストレージ装置の正D K C側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

【図11】本発明の実施の形態2によるストレージ装置の正D K C側の非同期周期処理の動作を示すフローチャートである。

40

【図12】本発明の実施の形態2によるストレージ装置のホストコンピュータからの同期ポイント指示の動作を示すフローチャートである。

【図13】本発明の実施の形態2によるストレージ装置の副D K C側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

【図14】本発明の実施の形態3によるストレージ装置のライトデータの転送動作を説明するための説明図である。

【図15】本発明の実施の形態3によるストレージ装置の正D K C側のライトデータをキャッシュメモリに格納する動作およびデータ転送動作を示すフローチャートである。

【図16】本発明の実施の形態3によるストレージ装置の副D K C側のライトデータをキ

50

キャッシュメモリに格納する動作を示すフローチャートである。

【図 17】本発明の実施の形態 5 によるストレージ装置のモード切替の動作を説明するための説明図である。

【図 18】本発明の実施の形態 5 によるストレージ装置の正 D K C 側のモード切替の動作を示すフローチャートである。

【図 19】本発明の実施の形態 5 によるストレージ装置の副 D K C 側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

【図 20】本発明の実施の形態 6 によるストレージ装置のモード切替の動作を説明するための説明図である。

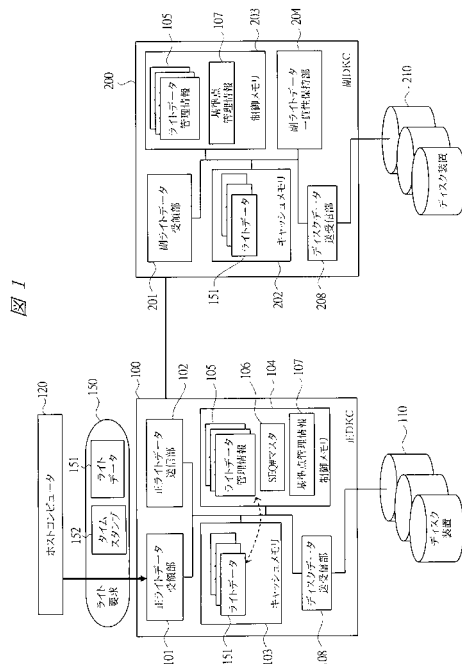
【図 21】本発明の実施の形態 6 によるストレージ装置の副 D K C 側のライトデータをキャッシュメモリに格納する動作を示すフローチャートである。

【符号の説明】

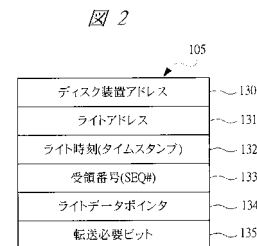
【 0 1 3 3 】

1 0 0 ... 正ディスク制御装置 (正 D K C)、1 0 1 ... 正ライトデータ受領部、1 0 2 ... 正ライトデータ送信部、1 0 3 ... キャッシュメモリ、1 0 4 ... 制御メモリ、1 5 1 ... ライトデータ、1 0 5 ... ライトデータ管理情報、1 0 6 ... S E Q # マスタ、1 0 7 ... 基準点管理情報、1 0 8 ... ディスクデータ送受信部、1 1 0 ... ディスク装置、1 3 0 ... ディスク装置アドレス、1 3 1 ... ライトアドレス、1 3 2 ... ライト時刻、1 3 3 ... 受領番号、1 3 4 ... ライトデータポインタ、1 3 5 ... 転送必要ビット、2 0 0 ... 副ディスク制御装置 (副 D K C)、2 0 1 ... 副ライトデータ受領部、2 0 2 ... キャッシュメモリ、2 0 3 ... 制御メモリ、2 0 4 ... 副ライトデータ一貫性保持部、2 0 8 ... ディスクデータ送受信部、2 1 0 ... ディスク装置。

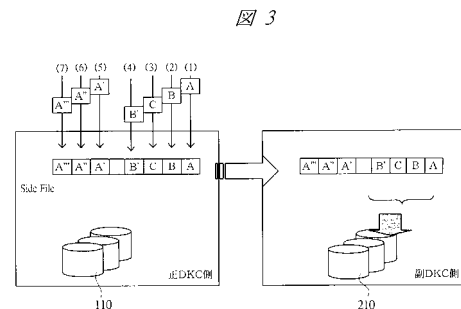
【図 1】



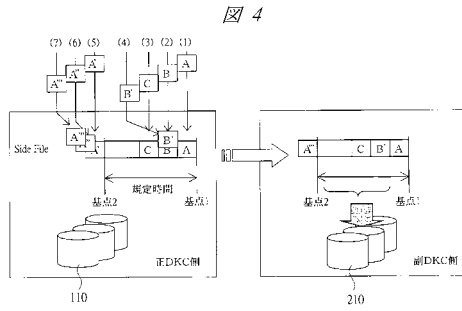
【図 2】



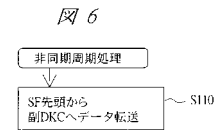
【図 3】



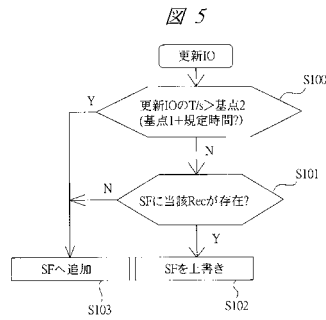
【図 4】



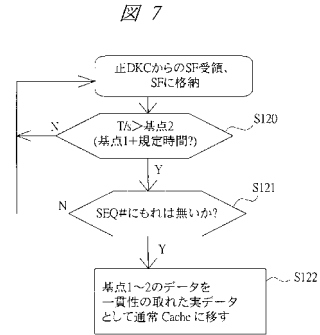
【図 6】



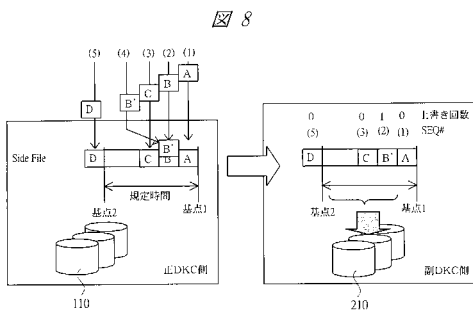
【図 5】



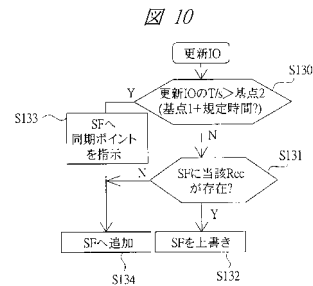
【図 7】



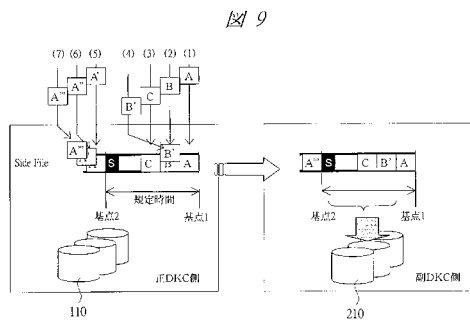
【図 8】



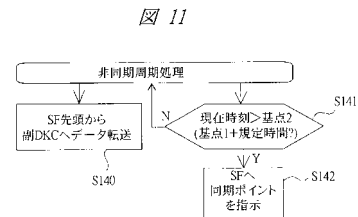
【図 10】



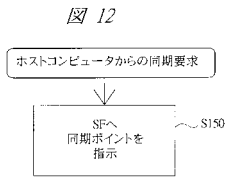
【図 9】



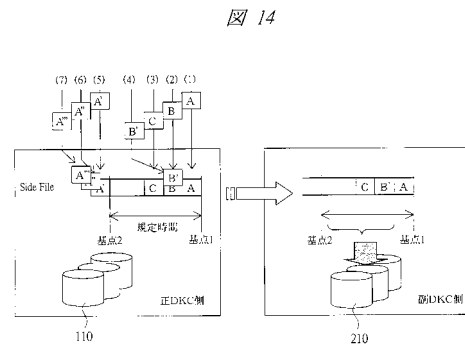
【図 11】



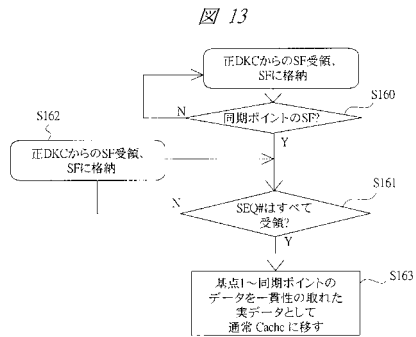
【 図 1 2 】



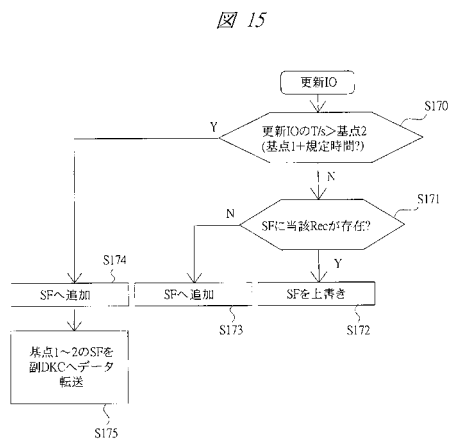
【 図 1 4 】



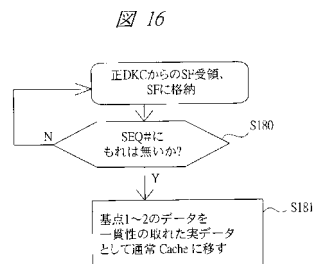
【 図 1 3 】



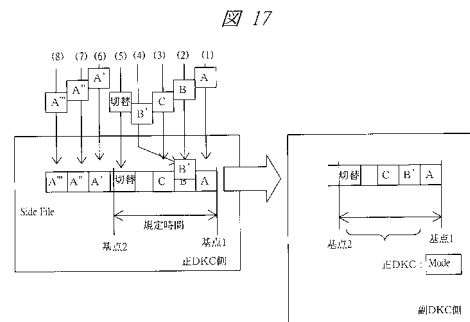
【 図 1 5 】



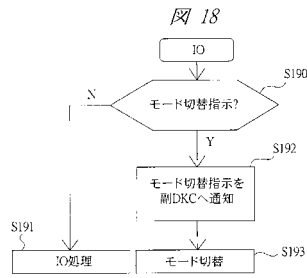
【 図 1 6 】



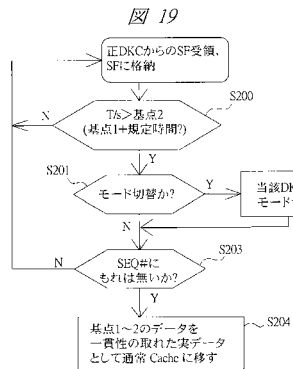
【 図 1 7 】



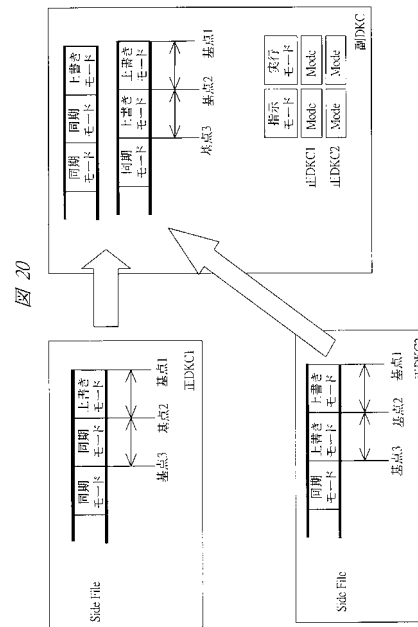
【図 18】



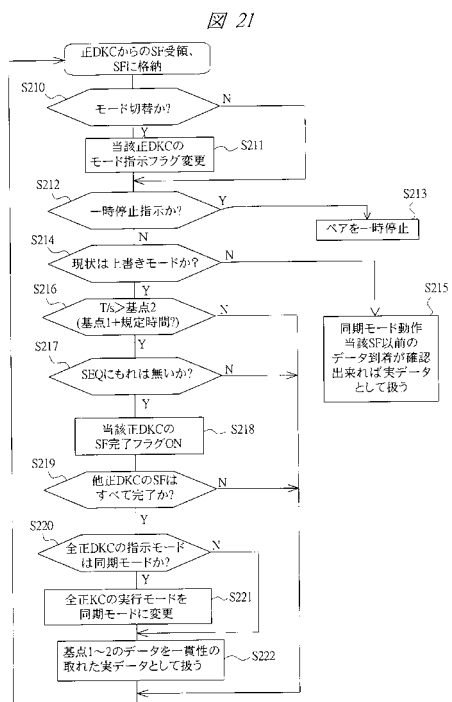
【図 19】



【図 20】



【図 21】



フロントページの続き

(56)参考文献 特開2002-318718(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06

G06F 12/00