

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2009-15476

(P2009-15476A)

(43) 公開日 平成21年1月22日(2009.1.22)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G06F 12/00 (2006.01)</b>	G06F 12/00 531R	5B065
<b>G06F 3/06 (2006.01)</b>	G06F 3/06 301M	5B082
	G06F 12/00 531J	

審査請求 未請求 請求項の数 15 O L (全 25 頁)

(21) 出願番号	特願2007-174717 (P2007-174717)	(71) 出願人	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22) 出願日	平成19年7月3日(2007.7.3)	(74) 代理人	100100310 弁理士 井上 学
		(74) 代理人	100098660 弁理士 戸田 裕二
		(72) 発明者	市川直子 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	岡田渡 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】 CDPリモート構成におけるジャーナル管理方法

(57) 【要約】

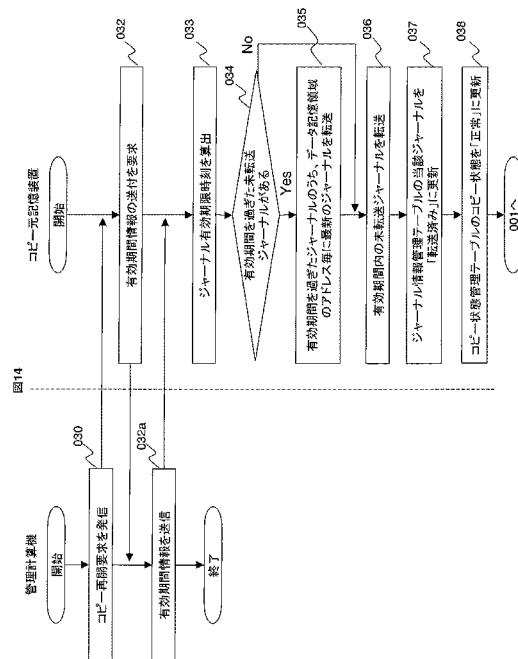
【課題】

転送元と転送先の記憶装置間でデータ転送が一時中断し、再開する場合に、転送先の記憶装置で必要とされる時間の範囲において任意の時点のデータを復旧することを目的とする。

【解決手段】

上記課題を解決するため、本発明の一形態では、第一のストレージシステムは、第二のストレージシステムにおけるデータの更新(書き込み)単位での書き込みデータの保持期間に従って、転送すべきジャーナルを特定し、第二のストレージシステムに転送する。保持期間を経過したデータは、アドレスごとに書き込みデータを選択して転送する。

【選択図】 図14



## 【特許請求の範囲】

## 【請求項 1】

計算機システムであって、

計算機がネットワークを介して利用する第一の記憶領域と前記第一の記憶領域のいずれか一のアドレスを更新する書込み要求に含まれる書込みデータを格納する第一のジャーナル記憶領域とを有する第一のストレージシステムと、

前記第一の記憶領域の複製データを格納する第二の記憶領域と第二のジャーナル記憶領域とを有する第二のストレージシステムと、を備え、

前記第一のストレージシステムは、

前記第一のジャーナル記憶領域に格納される前記第一の記憶領域に対する書込みデータの属性と、前記第二のジャーナル記憶領域の属性とを比較し、前記書込みデータを前記第二のストレージシステムに送信すべきか否かを判断する判断手段を有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、

前記書込みデータの属性には、時刻を特定する情報が含まれ、前記第二のジャーナル記憶領域の属性には、前記第二のジャーナル記憶領域に格納されるデータの保持期間が含まれ、

前記第一のストレージシステムは、さらに、書込みデータを前記第一のストレージシステムから前記第二のストレージシステムに転送する転送手段を備え、

20

前記判断手段は、前記第一のストレージシステムから前記第二のストレージシステムへ、書込みデータの転送を開始する場合に、前記時刻を特定する情報と前記保持期間と、を比較し、前記転送手段は、前記比較結果により前記保持期間内の書込みデータを前記第二のストレージシステムに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 2 記載の計算機システムであって、

前記判断手段は、さらに、前記比較結果、前記保持期間外の書込みデータのうち、同じアドレスに対する書込みデータが複数ある場合、前記書込みデータの属性に含まれる時刻を特定する情報により示す時刻が遅い書込みデータを特定し、

前記転送手段は、前記特定された書込みデータを前記第二のストレージシステムに転送することを特徴とする計算機システム。

30

## 【請求項 4】

請求項 3 記載の計算機システムであって、前記保持期間外で、前記特定された書込みデータ以外の前記同じアドレスに対する書込みデータを削除する削除手段を有することを特徴とする計算機システム。

## 【請求項 5】

請求項 4 記載の計算機システムであって、前記削除手段は、前記第一のジャーナル記憶領域に格納されるデータの保持期間の経過に応じて削除することを特徴とする計算機システム。

## 【請求項 6】

請求項 2 ないし 5 記載の計算機システムであって、

前記第一のストレージシステムは、前記書込みデータの属性を格納する属性格納記憶領域を有することを特徴とする計算機システム。

40

## 【請求項 7】

請求項 1 ないし 6 記載の計算機システムであって、

前記第一のストレージシステムは、

前記第二のジャーナル記憶領域の属性を示す属性情報を取得する属性情報取得手段を有することを特徴とする計算機システム。

## 【請求項 8】

請求項 2 ないし 7 記載の計算機システムであって、

50

前記第二のジャーナル記憶領域に格納される書込みデータの保持期間は、前記第一のジャーナル記憶領域に格納される書込みデータの保持期間より短い期間であることを特徴とする計算機システム。

【請求項 9】

請求項 2 ないし 8 記載の計算機システムであって、第二のストレージシステムは、第三の記憶領域を備え、保持期間外の複数のデータを用いてスナップショットを取得し、前記用いたデータの属性に含まれる時刻を特定する情報と関連付けて第三の記憶領域に格納するスナップショット取得手段を有することを特徴とする計算機システム。

【請求項 10】

第一の記憶領域と第一のジャーナル記憶領域とを有する第一のストレージシステムから第二の記憶領域と第二のジャーナル記憶領域とを有する第二のストレージシステムにデータを転送するリモートコピー制御方法であって、

計算機から受信するデータを用いて前記第一の記憶領域を更新し、前記データを前記第一の記憶領域の更新すべきアドレスと前記アドレスに対する書込み順序とに対応付けて前記第一のジャーナル記憶領域に格納し、前記第一のジャーナル記憶領域に格納されているデータのうち、前記第二のジャーナル記憶領域に格納されるデータの保持期間内のデータを前記第二のストレージシステムに送信し、

前記送信されるデータを前記第二のジャーナル記憶領域に格納し、前記第二のジャーナル記憶領域に格納されているデータを用いて前記第二の記憶領域を更新する、ことを特徴とするリモートコピー制御方法。

【請求項 11】

請求項 10 記載のリモートコピー制御方法であって、さらに、前記第一のジャーナル記憶領域に格納されているデータのうち、前記アドレスごとに前記保持期間外のいずれかのデータを前記第二のストレージシステムに送信することを特徴とするリモートコピー制御方法。

【請求項 12】

請求項 11 記載のリモートコピー制御方法であって、前記保持期間外のいずれかのデータは、前記第一のジャーナル記憶領域に格納されている前記保持期間外のデータのうち、前記アドレスごとに前記書込み順序がもっとも遅いものであることを特徴とするリモートコピー制御方法。

【請求項 13】

請求項 10 ないし 12 記載のリモートコピー制御方法であって、前記第一のストレージシステムから前記第二のストレージシステムへのデータ転送処理を再開する場合、前記第一のジャーナル記憶領域に格納されているデータが、前記第二のジャーナル記憶領域に格納されるデータの保持期間内か否かを判断することを特徴とするリモートコピー制御方法。

【請求項 14】

ストレージシステムであって、計算機と他のストレージシステムとにネットワークを介して接続されるネットワークインターフェースと、

前記ネットワークに接続される制御部と、前記制御部に接続される複数のディスク装置と、を備え、前記いずれか以上のディスク装置は、前記計算機が利用する第一の記憶領域と、前記計算機から受信する前記第一の記憶領域に対する書込みデータを格納する第二の記憶領域とのうち少なくとも一方を構成し、

前記制御部は、前記書込みデータごとに前記他のストレージシステムに前記ネットワークインターフェースを介して送信するか否かを、前記書込みデータごとに対応付けられている所定の時刻と、前記他のストレージシステムでの書込みデータの保持期限とに基づいて

10

20

30

40

50

判断する、ことを特徴とするストレージシステム。

【請求項 15】

請求項 14 記載のストレージシステムであって、

前記他のストレージシステムにおける前記書込みデータごとにデータを保持する保持期限は、前記書込みデータごとにデータが格納される前記他のストレージシステムが有する第三の記憶領域の属性に含まれ、

前記制御部は、前記第三の記憶領域の属性を取得することを特徴とするストレージシステム。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明で開示される技術は、計算機と記憶装置とを有する計算機システムに関し、特に、記憶装置間のデータコピー制御に関する。

【背景技術】

【0002】

データの更新履歴を保存し、データの更新（書き込み）単位で任意の時刻のデータを復元する技術がある。特許文献 1 では、データの更新履歴情報（以下、ジャーナル）及び制御情報を、第一の記憶装置から、距離を隔てた場所に設置した第二の記憶装置に転送することで、復元可能な範囲において任意の時点のデータを第一の記憶装置または第二の記憶装置において復元する技術が開示されている。また、特許文献 2 においては、前述のジャーナル転送時に、同一アドレスへの書き込みデータについては、最新のジャーナルを上書き更新することにより、ジャーナルデータを圧縮し、データ転送量を小容量化する技術が開示されている。

20

【0003】

【特許文献 1】特開 2005 - 222110 号公報

【特許文献 2】米国出願公開公報 US 20050028022 A 1

【発明の開示】

【発明が解決しようとする課題】

【0004】

第一及び第二の記憶装置において、任意時点のデータを復元可能とする技術を適用する際、記憶装置間のデータ転送が一定時間中断されたとしても、データ転送再開後に、データ転送中断前と同様に、任意時点のデータを復元可能であることが望ましい。

30

【0005】

特許文献 1 では、ジャーナルの転送処理が障害や計画停止などによって一定期間停止すると、転送が再開され第二のストレージ装置に転送完了するまで、第一の記憶装置にジャーナルは蓄積される。そして、ジャーナルの転送が停止していた期間に第一の記憶装置に蓄積したジャーナルを、ジャーナルの転送を再開する際に全て転送するため、ジャーナル転送量は停止期間の長さに応じて肥大化し、転送に要する時間が長期化し、第二の記憶装置での運用再開が遅れる場合がある。一方、ジャーナルの転送に要する時間を短くするために、特許文献 2 のようにと、ジャーナルデータを圧縮すると、同一アドレスへの書き込みに対する更新履歴情報は失われ、データの更新（書き込み）単位で任意の時刻のデータを復元することができず、第二の記憶装置で要求される業務を復旧条件が満足できない場合がある。

40

【課題を解決するための手段】

【0006】

上記課題を解決するため、本発明の一形態では、第一のストレージシステムは、第二のストレージシステムにおけるデータの更新（書き込み）単位での書き込みデータの保持期間に従って、転送すべきジャーナルを特定し、第二のストレージシステムに転送する。保持期間を経過したデータは、アドレスごとに書き込みデータを選択して転送する。

【発明の効果】

50

## 【 0 0 0 7 】

上述した本発明の一形態により、転送先のストレージシステムで必要とされる時間の範囲において任意の時点のデータを復旧する。

## 【 発明を実施するための最良の形態 】

## 【 0 0 0 8 】

以下、本発明の実施の形態を、図面を参照しながら説明する。

## 【 実施例 1 】

## 【 0 0 0 9 】

以下に本発明の第一の実施例について図 1 ~ 図 1 8 を用いて説明する。図 1 は、実施例 1 の構成を模式的に示した図である。

10

## 【 0 0 1 0 】

図 1 において、ホスト計算機 4 0 0 と記憶装置（ストレージシステム）1 0 0 は、データ I/O 用ネットワーク 1 0 1 を介して接続されている。また、記憶装置 1 0 0 と記憶装置（ストレージシステム）2 0 0 はデータ転送用ネットワーク 1 0 3 を介して接続され、相互に通信する。データ I/O 用ネットワーク 1 0 1 とデータ転送用ネットワーク 1 0 3 は、ファイバチャネルや IP ネットワーク等の一般的なネットワーク接続形態から成る。

## 【 0 0 1 1 】

また、本実施形態は記憶装置 1 0 0 および記憶装置 2 0 0 のデータ通信や記憶領域の割当や解除の構成に関する管理を行う管理計算機 3 0 0 を有する。管理計算機 3 0 0 は、管理用ネットワーク 1 0 2 を介して記憶装置 1 0 0、記憶装置 2 0 0 に接続される。管理用ネットワーク 1 0 2 は IP ネットワーク等の一般的なネットワーク接続装置で構成される。また、この管理用ネットワーク 1 0 2 は、前述のデータ I/O 用ネットワーク 1 0 1 やデータ転送用ネットワーク 1 0 3 と同一のものを共有する形態であってもよい。管理計算機 3 0 0 と記憶装置 1 0 0 及び記憶装置 2 0 0 はこの管理用ネットワーク 1 0 2 を介して相互に管理情報の送受信を行う。また、管理計算機 3 0 0 は、管理ネットワーク 1 0 2 を介して計算機 4 0 0、5 0 0 とも接続し、記憶装置 1 0 0 を利用するに必要な情報の送受信を行なう。

20

## 【 0 0 1 2 】

記憶装置 1 0 0 は、記憶領域として、データ記憶領域 1 2 0、ジャーナルデータ記憶領域 1 2 1、ジャーナルメタデータ記憶領域 1 2 2 及びスナップショット記憶領域 1 2 3 を有する。データ記憶領域 1 2 0 は、ホスト計算機 4 0 0 からの書き込み要求を受信した際に、指定されたアドレスに書き込み要求に含まれるデータを用いて更新され、その後、計算機に利用される。また、ジャーナルデータ記憶領域 1 2 1 は、後述するジャーナル取得プログラム 1 3 3 によって取得された、データ記憶領域 1 2 0 に対する更新データ（以下、ジャーナルデータ）が格納される記憶領域である。ジャーナルメタデータ記憶領域 1 2 2 は、データ記憶領域 1 2 0 に対する更新履歴情報（ジャーナルメタデータ）を格納する記憶領域である。ただし、記憶装置によっては、ジャーナルデータ記憶領域 1 2 1 とジャーナルメタデータ記憶領域 1 2 2 とは同一の記憶領域であってもよい。

30

## 【 0 0 1 3 】

さらに、スナップショット記憶領域 1 2 3 は、後述するスナップショット取得プログラムによって作成されたスナップショットを格納する記憶領域である。“スナップショット”とは、本明細書において、ある指定時点における一つ又は複数のデータ記憶領域 1 2 0 の全体又は部分的なデータイメージを意味するものとする。

40

## 【 0 0 1 4 】

図 2 は、記憶装置 1 0 0 のハードウェア構成を示す。記憶装置 1 0 0 は、ストレージコントローラ 1 6 0、ハードディスク 1 1 0、プログラムメモリ 1 3 0、キャッシュメモリ 1 4 0、CPU 1 5 0 を有する。また、記憶装置外部との通信を行なうために、用途別にストレージコントローラ 1 6 0 に接続された I/O 用通信インタフェース 1 7 0、管理用通信インタフェース 1 8 0 及びデータ転送用通信インタフェース 1 9 0 を有する。

## 【 0 0 1 5 】

50

キャッシュメモリ 140 は、物理的には一般的な半導体記憶装置であればよく、汎用計算機のそれと同様に一時的なデータ記憶領域として利用する。

【0016】

ハードディスク 110 は、例えば 1 台以上の磁気ディスク装置、すなわち一般的にハードディスクと呼ばれる記憶デバイスで、論理的には複数の記憶領域を構成し、データ記憶領域 120、ジャーナルデータ記憶領域 121、ジャーナルメタデータ記憶領域 122、スナップショット記憶領域 123 のいずれか一以上を構成する。

【0017】

プログラムメモリ 130 は、物理的には磁気ディスク装置あるいは半導体記憶装置で構成する記憶領域である。記憶装置の動作を担う各種プログラム群および情報を保持し、制御部または CPU が各種プログラム群や情報を読み出すことにより各種プログラムを実行する。プログラムメモリ 130 は管理情報入出力プログラム 131、ジャーナル取得プログラム 132、データコピープログラム 133、データ復元プログラム 134、ペア構成管理テーブル 135、コピー状態管理テーブル 136、ジャーナル情報管理テーブル 137、スナップショット取得プログラム 138、スナップショット管理テーブル 139 を格納する。

10

【0018】

管理情報入出力プログラム 131 は、記憶装置 100 と管理計算機 300 及び記憶装置 200 との間で管理情報の送受信を行う。また、受信した管理情報をプログラムメモリ 130 内のプログラム又はテーブルへ伝達する。例えば、管理計算機 300 からデータコピー再開要求が発信された場合、その情報を受信し、データコピープログラム 132 へ伝達する。

20

【0019】

ジャーナル取得プログラム 132 は、ホスト計算機 400 からデータ記憶領域 120 に発信された全ての書き込み動作に対してジャーナルデータと、ジャーナルメタデータを採取する。ジャーナルメタデータは、少なくとも、ジャーナルデータが書き込まれた時刻、書き込み先データ記憶領域のアドレス情報、書き込み先データ記憶領域内のアドレス（書き込み開始位置）、書き込み順序番号を含む。書き込み順序番号は、データが書き込まれた順にジャーナル取得プログラム 132 が付与する番号である。ジャーナル取得プログラム 132 は採取したジャーナルデータ及びジャーナルメタデータをジャーナル記憶領域 121 及びジャーナルメタデータ記憶領域 122 にそれぞれ格納する。尚、本明細書においては、ジャーナルデータとジャーナルメタデータを合わせて単に“ジャーナル”と称する場合がある。

30

【0020】

データコピープログラム 133 は、データ記憶領域 120 へのデータ書き込み処理とは非同期に、ジャーナルデータ及びジャーナルメタデータを、後述するペア構成管理テーブル 135 においてペア（対となる組み合わせ）として定義されたデータ記憶領域 120 を有する記憶装置 200 へ転送する。また、データコピープログラム 133 は、データコピー処理の状態に関する情報を蓄積するコピー状態管理テーブル 136 を必要に応じて更新する。例えば、データ転送用ネットワークの障害などによってデータコピー処理が異常終了となった場合には、コピー処理を中断し、当該ペアのコピー状態管理テーブル 136 が管理するコピー状態情報を「停止」に更新する。コピー状態管理テーブル 136 の詳細については後述する。ただし、データコピープログラムは、障害が発生してからコピー状態情報を更新するまでに、一定期間の間、コピー処理を再実行（リトライ）してもよい。

40

【0021】

また、データコピープログラム 133 は、コピー状態が「停止」である状態から、データコピー処理を再開する際、管理計算機 300 から取得したジャーナル有効期間に関する情報に基づき、記憶装置 200 に転送するジャーナルを選択する。ジャーナル有効期間とは、管理計算機 300 によって管理される、各データ記憶領域に対応付けられるジャーナルの保持期間に関する情報である。

50

## 【 0 0 2 2 】

図 3 は、データコピー処理再開時に記憶装置 2 0 0 へ転送するジャーナルの選択方法を示す。図 3 は、時間軸上の各時刻において、あるデータ記憶領域 1 2 0 内のアドレスに対するデータ書き込み処理に対して取得されたジャーナルと、データコピー再開時に転送するジャーナルの関係さらに、ジャーナルの有効期間との関係を模式的に表す。時間軸上に配置した黒点はそれぞれの時刻にジャーナルが取得されたことを意味し、黒点に付与した数値は、ジャーナルメタデータ毎に付与された書き込み順序番号である。そして、記憶装置 1 0 0 が有するコピーペア 0 1 に関連付けられる記憶領域のジャーナル有効期間が 24 時間であり、記憶装置 2 0 0 が有するコピーペア 0 1 に関連付けられる記憶領域のジャーナル有効期間が 1 0 時間であることを矢印で示す。

10

## 【 0 0 2 3 】

図 3 の例に示すように、記憶装置 2 0 0 におけるジャーナル有効期間が 1 0 時間であって、時刻 2 0 : 0 0 に記憶装置 1 0 0 から記憶装置 2 0 0 にデータコピーを再開する場合、1 0 時間前である時刻 1 0 : 0 0 以降に取得した全てのジャーナル、すなわち、書き込み順序番号 4、5、6 で表されるジャーナルと、時刻 1 0 : 0 0 以前に取得したジャーナルのうち、最新のもの、すなわち、書き込み順序番号 3 で表されるジャーナルを、書き込み順序番号の順に記憶装置 2 0 0 に転送する。

## 【 0 0 2 4 】

一方、データコピー先となる記憶装置 2 0 0 においても、記憶装置 1 0 0 と同様のデータコピープログラム 1 3 3 及び後述のコピー構成管理テーブルを有する。記憶装置 2 0 0 において、データコピープログラム 1 3 3 は、記憶装置 1 0 0 から受領したジャーナルデータ及びジャーナルメタデータを、自装置が有するジャーナルデータ記憶領域 1 2 1 及びジャーナルメタデータ記憶領域 1 2 2 にそれぞれ格納する。また、ジャーナルメタデータに付与された書き込み順序番号の順に、当該データ記憶領域 1 2 0 に書き込み処理を実施する。

20

## 【 0 0 2 5 】

スナップショット取得プログラム 1 3 8 は、データ記憶領域 1 2 0 の指定時点でのスナップショットを取得する。“スナップショット”とは、本明細書において、ある指定時点における一つ又は複数のデータ記憶領域の全体又は部分的なデータイメージを意味するものとする。取得されたスナップショットは、スナップショット記憶領域 1 2 3 に格納される。

30

## 【 0 0 2 6 】

データ復元プログラム 1 3 4 は、管理計算機 3 0 0 から受領したデータ復元指定時刻におけるデータを復元する。データ復元プログラム 1 3 4 によって実行されるデータ復元方法を、図 4 を用いて説明する。尚、本節によって説明するデータ復元技術は特許文献 1 において公知である。

## 【 0 0 2 7 】

図 4 は、過去の時点の記憶領域 1 2 0 に対するスナップショットと、時間軸上の各時刻において取得されたジャーナルとの関係を模式的に表す。時間軸上に配置した黒点はそれぞれの時刻にジャーナルが取得されたことを意味し、黒点に付与した数値は、ジャーナルメタデータ毎に付与された書き込み順序番号である。

40

## 【 0 0 2 8 】

データ復元プログラム 1 3 4 は、管理計算機 3 0 0 からデータ復旧指定時刻（例えば、時刻 2 0 : 0 0 とする）を受領すると、スナップショットが表す時刻 0 6 : 0 0 とデータ復旧目標時刻 2 0 : 0 0 との間を取得されたジャーナルデータを、ジャーナルメタデータに割り当てられた書き込み順序番号の順に、スナップショットデータに適用する。

## 【 0 0 2 9 】

例えば、図 4 において、時刻 0 6 : 0 0 と 2 0 : 0 0 の間を取得されたジャーナルは書き込み順序番号 1、2、3、4 を付加されたジャーナルであり、書き込み順序番号 5 を付加されたジャーナルは対象外となる。従って、データ復元プログラム 1 3 4 は、スナップ

50

ショットにこれらのジャーナルを順に適用することにより、データ復旧指定時刻におけるデータを復元することができる。

【0030】

ペア構成管理テーブル135は、データコピープログラム133がデータコピー処理を実施する際の、記憶装置100及び記憶装置200内のデータ記憶領域のペア（対となる組み合わせ）構成が定義されている。ペア構成とは、データコピープログラムによるデータコピー処理のコピー元とコピー先となるデータ記憶領域120間の対応付けを意味する。ペア情報管理テーブル135の例を図5に示す。図5において、ペア情報管理テーブル135は、コピーペア識別子1350、コピー元記憶装置識別子1351、コピー元データ記憶領域識別子1352、コピー先記憶装置識別子1353、コピー先データ記憶領域識別子1354のカラムを有する。

10

【0031】

コピーペア識別子1350は、コピーペアを特定する識別子である。コピー元データ記憶領域識別子1352及びコピー元記憶装置識別子1351は、それぞれ、コピーペア識別子1350により特定されるコピーペアを構成するデータ記憶領域のうち、コピー元のデータ記憶領域を特定する識別子とコピー元のデータ記憶領域を有する記憶装置を特定する識別子である。コピー先データ記憶領域1354及びコピー先記憶装置識別子1353は、コピーペア識別子1350により構成されるコピー先データ記憶領域を特定する識別子及びコピー先データ記憶領域を有する記憶装置を特定する識別子である。

【0032】

例えば、コピーペア識別子01で表されるコピーペアは、記憶装置識別子1000により特定される記憶装置内に存在するデータ記憶領域識別子00:01により特定されるデータ記憶領域に対する書き込みデータが、データコピープログラム133によって、記憶装置識別子2000により特定される記憶装置が存在するデータ記憶領域識別子0A:01により特定されるデータ記憶領域にコピーされる関係である。

20

【0033】

コピー状態管理テーブル136は、データコピープログラム133によるコピーペアごとにデータコピー処理の状態に関する管理情報を蓄積する。コピー状態管理テーブル136の例を図6に示す。コピー状態管理テーブル136は、少なくともコピーペア識別子1360、コピー処理の状態1361、コピー状態更新時刻1362などの情報を保持する。コピー処理の状態1361は、コピーペア識別子1360により特定されるコピーペアを構成するデータ記憶領域間のコピー状態を示す。コピー処理の状態が「正常」の場合は、コピーペア間の通信障害を検知していない状態で、ジャーナルを転送可能な状態であることを示す。コピー処理の状態が「停止」の場合は、通信障害やシステム運用操作などの原因により、ジャーナルが転送できず、記憶装置100のジャーナルデータ記憶領域に転送できないジャーナルを蓄積している状態であることを示す。コピー状態更新時刻は、コピーペアのコピー状態を更新した時刻を示す。

30

【0034】

例えば、図6において、識別子01のコピーペアは時刻09:30:00に「停止」に更新されたことを表している。

40

【0035】

図7は、ジャーナル情報管理テーブル137の例である。ジャーナル情報管理テーブル137は、ジャーナル取得プログラム132が取得するジャーナルに関する管理情報を保持する。ジャーナル管理情報テーブル137は、少なくとも書き込み順序番号1370、ジャーナル取得時刻1371、データ記憶領域識別子1372、データ記憶領域データ開始位置1373、ジャーナルデータ記憶領域識別子1374、転送状況1375に関する情報を蓄積する領域を有する。

【0036】

書き込み順序番号1370は、ジャーナルデータ記憶領域に保持されるジャーナルのデータ記憶領域に対する書き込み順序を表し、同順序に従ってそれぞれのジャーナルデータに含

50



まれるデータがデータ記憶領域に書き込まれたことを表す。ジャーナル取得時刻 1 3 7 1 は、データ記憶領域に書き込みが発生し、同書き込みデータのジャーナルを取得した時刻を表す。データ記憶領域識別子 1 3 7 2 は、書き込みが発生したデータ記憶領域の識別情報を表し、ジャーナル取得プログラム 1 3 2 は、同データ記憶領域への書き込みをジャーナルとして記録する。

**【 0 0 3 7 】**

データ記憶領域内のアドレス 1 3 7 3 は、同ジャーナルが該当する、書き込みが発生したデータ記憶領域内の位置情報を表す。ジャーナル記憶領域識別子 1 3 7 4 は、同ジャーナルデータが記録されたジャーナルデータ記憶領域の識別情報を表す。転送状況 1 3 7 5 は、同ジャーナルデータがコピー元記憶装置からコピー先記憶装置に転送されたかどうかを表す情報であり、転送されていれば“転送済み”、転送されていなければ“未転送”といった文字列で表現する。また、転送が不要であることを示す“転送不要”といった文字列で表現する。本実施例では転送状況を文字列で表現したが、“0”または“1”などの真偽値で表現してもよい。

10

**【 0 0 3 8 】**

例えば、図 7 において、書き込み順序番号が 0 1 で表されるジャーナルは、時刻 0 8 : 0 8 : 0 0 に取得された、識別子 0 0 : 0 1 で表されるデータ記憶領域内のアドレス F 5 の位置に書き込まれたデータに対するジャーナルであり、更新データであるジャーナルデータは記憶領域 0 B : 0 1 に格納されており、データコピープログラム 1 3 3 によって既に記憶装置 2 0 0 へ転送済みであることを表す。

20

**【 0 0 3 9 】**

図 8 は、スナップショット管理テーブルの例を示す。スナップショット管理テーブル 1 3 9 は、記憶装置 1 0 0 が保持する、データ記憶領域 1 2 0 のスナップショットに関する情報を管理する。スナップショット管理テーブル 1 3 9 は、自記憶装置においてスナップショットを取得した時刻情報（フィールド 1 3 9 0）、取得したスナップショットの格納場所となるスナップショット記憶領域識別子（1 3 9 1）、スナップショットを作成した元データに相当するデータ記憶領域の識別子（フィールド 1 3 9 2）、同データ記憶領域への書き込みのジャーナルデータが格納されたジャーナルデータ記憶領域の識別子（フィールド 1 3 9 3）、同データ記憶領域への書き込みのジャーナルメタデータが格納されたジャーナルメタデータ記憶領域の識別子（フィールド 1 3 9 4）により構成する。

30

**【 0 0 4 0 】**

例えば、図 8 において、時刻情報 0 9 : 0 0 に相当するスナップショットは、1 0 : 0 1 で表されるスナップショット記憶領域に格納される。また、同スナップショットはデータ記憶領域 0 0 : 0 1 に対して作成されたものであり、これに関するジャーナルデータ及びジャーナルメタデータは、それぞれ識別子 0 B : 0 1、0 C : 0 1 で表される記憶領域に格納される。

**【 0 0 4 1 】**

記憶装置 2 0 0 の構成は、記憶装置 1 0 0 と同様のものでよい。また、記憶装置 2 0 0 内のプログラムメモリ 1 3 0 においては、ジャーナル取得プログラム 1 3 2、コピー状態管理テーブル 1 3 6 は必ずしも必要ではない。ただし、運用によっては記憶装置 1 0 0 と記憶装置 2 0 0 が同様の手順に従い、逆方向にデータ転送を行ってもよく、その場合には上記プログラムを備えていればよい。

40

**【 0 0 4 2 】**

図 9 は、管理計算機 3 0 0 の構成を示す。管理計算機 3 0 0 は、プログラムメモリ 3 2 0、出力装置 3 3 0、入力装置 3 4 0、管理用通信インタフェース 3 5 0、CPU 3 6 0、キャッシュメモリ 3 7 0 を有し、それぞれがバス接続されている。管理計算機 3 0 0 のハードウェア構成は、例えば汎用計算機（PC）のそれと同様であればよい。例えば入力装置 3 5 0 はキーボードやマウスといったデバイスであればよく、出力装置 3 4 0 は CRT や LCD などのディスプレイ装置とビデオ出力デバイスであればよい。同様に管理用通信インタフェース 3 6 0 はイーサネット（登録商標）などの汎用的な通信デバイスであれば

50

よい。プログラムメモリ320は磁気記憶装置もしくは半導体記憶装置によるデータ記憶デバイスであればよい。プログラムメモリ320は、少なくとも、管理情報入出力プログラム321、ジャーナル有効期間管理テーブル322を格納する。プログラムメモリ320に格納されるプログラムは、CPU360が読み出し、実行する。また、CPU360が各種プログラムを実行する際は、プログラムメモリに格納されている必要テーブルを参照する。

#### 【0043】

管理情報入出力プログラム321は、管理計算機300と記憶装置100、記憶装置200との間で管理情報の送受信を行う。また、受信した管理情報をプログラムメモリ320内のプログラム又はテーブルへ伝達する。つまり、CPU360が、管理情報入出力プログラム321の実行することにより受信する管理情報をプログラムメモリに格納したり、管理情報を用いて別のプログラムを実行する。

10

#### 【0044】

図10は、ジャーナル有効期間管理テーブルの例を示す。ジャーナル有効期間管理テーブル322は、記憶装置100及び記憶装置200に格納されたジャーナルの有効期間に関する情報を保持する。ジャーナル有効期間管理テーブル322は、少なくとも、コピーペア識別子(フィールド3220)、正サイトにおけるジャーナル有効期間(フィールド3221)、副サイトにおけるジャーナル有効期間(フィールド3222)を蓄積する。コピーペア識別子は、記憶装置100内のペア構成管理テーブル135のそれと一致する。また、ジャーナル有効期間は、記憶装置100又は記憶装置200内のジャーナルを保持する期間を意味する。

20

#### 【0045】

例えば、コピーペア識別子01に関しては、正サイトにおけるジャーナル有効期間が24時間であり、副サイトにおけるジャーナル有効期間が10時間である。すなわち、正サイトでは、現在から24時間前までに更新されたデータに対するジャーナルを保管する。また、副サイトでは、現在から10時間前までに更新されたデータに対するジャーナルを格納する。

#### 【0046】

ジャーナル有効期間を過ぎたジャーナル、すなわち、上記の例であれば、正サイトで現時点から24時間前の時刻よりも前の更新に対するジャーナルは、本記憶システムにおいて格納する必要はなく、削除してもよい。あるいは、対応するデータ記憶領域120内のアドレス毎に最新のジャーナルのみを残して削除してもよい。あるいは、過去のスナップショットまたはその複製にジャーナルデータを書き込み、別の時点におけるスナップショットを作成し、保管してもよい。

30

#### 【0047】

本実施形態における処理を、図11、12、13、14、15を用いて解説する。

#### 【0048】

図11は、本実施形態における計算機システムの運用の流れを表す状態遷移図である。当該計算機システムにおいて、通常時はコピー状態管理テーブルの「コピー状態」が「正常」であり、図12を用いて後述する「正常」時の所定の処理を実施する(ステップ001)。コピー状態が正常のとき、データ転送用ネットワーク障害や運用管理操作など、コピー停止要因となるイベントが発生すると、コピー状態管理テーブルの「コピー状態」が「停止」に遷移する(ステップ003)。停止状態にあつては、図13-Aに後述する「停止」時の所定の処理を実施する。コピー状態が停止のとき、停止要因が取り除かれ、再開操作が行われると、「正常」状態に遷移して「正常」時の所定の処理を実施する。

40

#### 【0049】

図12は、「正常」時の処理フローを示す。

記憶装置100内のストレージコントローラ160が管理情報入出力プログラム131を実行し、ホスト計算機400からデータ記憶領域120への書き込み要求を受領する(ステップ011)。次に、ストレージコントローラは、ジャーナル取得プログラム132を

50

実行し、この書き込みデータに対するジャーナルデータ及びジャーナルメタデータを作成し、それぞれジャーナルデータ記憶領域 1 2 1 及びジャーナルメタデータ記憶領域 1 2 2 に格納する（ステップ 0 1 2）。

【 0 0 5 0 】

そして、ストレージコントローラ 1 6 0 は、データコピープログラム 1 3 3 を読み出して実行する。まず、ストレージコントローラ 1 6 0 は、コピー状態管理テーブル 1 3 6 を参照し、コピー状態が「正常」であれば、ジャーナルデータ及びジャーナルメタデータを、データ転送用通信インタフェース 1 0 3 を介してペア構成管理テーブル 1 3 5 においてペアであると定義された記憶領域 1 2 0 を有する記憶装置 2 0 0 に転送する（ステップ 0 1 3）。この間、記憶領域 1 2 0 へ書き込みデータは、データコピープログラム 1 3 3（ステップ 0 1 3）の処理とは非同期に、ストレージコントローラ 1 6 0 が一般的な書き込み処理と同様に処理する。

10

【 0 0 5 1 】

また、記憶装置 2 0 0 のストレージコントローラは、ジャーナルデータ及びジャーナルメタデータを受領すると、ジャーナル取得プログラムを読み出し、受領したジャーナルデータ及びジャーナルメタデータを。それぞれ自記憶装置内のジャーナルデータ記憶領域 1 2 1 及びジャーナルメタデータ記憶領域 1 2 2 に格納する（ステップ 0 1 4 a）。次に、記憶装置 2 0 0 のストレージコントローラは、書き込み順序番号の順にデータ記憶領域 1 2 0 にステップ 0 1 4 a でジャーナルデータ記憶領域に格納したジャーナルデータを書き込む処理を実施する（ステップ 0 1 4 b）。そして、記憶装置 2 0 0 のコントローラは、完了通知メッセージを記憶装置 1 0 0 に送信する（ステップ 0 1 4 c）。完了通知メッセージは、ステップ 0 1 4 a とステップ 0 1 4 b の実行結果が正常終了であればその旨を含み、ステップ 0 1 4 a 及び 0 1 4 b の実行結果が異常終了であればエラーが発生したこととエラーの種類とを示すエラー番号とを含む。

20

【 0 0 5 2 】

記憶装置 2 0 0 へのデータコピー処理が正常に終了すると（ステップ 0 1 5 の Yes）、データコピープログラム 1 3 3 を実行する記憶装置 1 0 0 のストレージコントローラ 1 6 0 は、ジャーナル情報管理テーブル 1 3 7 における当該ジャーナルの「転送状況」を「転送済み」に更新し（ステップ 0 1 6）、当該書き込みデータに対する処理は終了する。

【 0 0 5 3 】

また、データコピー処理がエラーとなった場合（ステップ 0 1 5 の No）、記憶装置 1 0 0 のストレージコントローラ 1 6 0 は、ペア構成管理テーブル 1 3 6 の「コピー状態」を「停止」に更新する（ステップ 0 1 7）。ただし、ストレージコントローラ 1 6 0 は「コピー状態」を更新するまでに一定期間の間、コピー処理を再実行（リトライ）してもよい。また、記憶装置 1 0 0 がステップ 0 1 3 で送信したコピー処理に対し、コピー先である記憶装置 2 0 0 が一定期間応答しなかった場合は、通信障害による異常終了などに見なし、ステップ 0 1 5 の判定結果を No としてもよい。ステップ 0 1 7 の結果、コピー状態が「停止」となることから、同システムは図 1 1 におけるステップ 0 0 3 の状態に遷移する。

30

【 0 0 5 4 】

図 1 2 がコピー処理の実行結果に基づく「停止」状態への遷移のケースだったのに対し、図 2 1 は運用管理操作によるコピー停止処理フローを表す。管理計算機 3 0 0 は、管理者の入力装置の操作により、コピー停止要求の入力を受ける。この入力を受け、管理計算機 3 0 0 はコピー停止要求メッセージを記憶装置 1 0 0 に送信する（ステップ 0 1 8）。記憶装置 1 0 0 は、コピーペア管理テーブルのコピー状態を「停止」に更新し（ステップ 0 1 7）、完了通知を管理計算機 3 0 0 に送信する（ステップ 0 1 9）。この結果、同システムは図 1 1 におけるステップ 0 0 3 の状態に遷移する。

40

【 0 0 5 5 】

図 1 3 は、コピー状態が「停止」である場合のジャーナル管理を示す。記憶装置 1 0 0 のストレージコントローラ 1 6 0 が、管理情報入出力プログラム 1 3 1 を読み出し実行す

50

る処理を説明する。ストレージコントローラ 160 は、ホスト計算機 400 からデータ記憶領域 120 への書き込み要求を受領する。(ステップ 021) ストレージコントローラ 160 は、ジャーナル取得プログラム 132 を実行し、この受領した書き込みデータに対するジャーナルメタデータを作成し、ジャーナルメタデータ記憶領域 121 に格納し、ジャーナルデータをジャーナルデータ記憶領域に格納する(ステップ 022)。この間、記憶領域 120 への書き込みデータは、データコピープログラム 133 を実行する処理とは非同期に、ストレージコントローラ 160 が一般的な書き込み処理と同様に処理する。

【0056】

図 14 は、コピー再開時の処理、すなわち、コピー状態が「停止」の状態から、管理者の差操作による入力を受け、コピー状態を「正常」に戻す際の処理の流れを示す。

10

【0057】

まず、管理者が管理計算機 300 の入力装置から、一つのコピーペアあるいは複数のコピーペアを指定するコピー再開要求を発信し、管理計算機 300 がコピー再開要求を記憶装置 100 に送信する(ステップ 030)。記憶装置 100 のストレージコントローラ 160 は、データコピープログラム 133 を実行し、管理計算機 300 が有するジャーナル有効期間管理テーブル 322 に記録されている、コピー再開要求に記載されたコピーペアの記憶装置 200 側におけるジャーナル有効期間の送付を要求する(ステップ 032)。

【0058】

要求を受けた管理計算機 300 は、同コピー先記憶領域の有効期間情報を送信する(ステップ 032 a)。

20

【0059】

データコピープログラム 133 を実行する記憶装置 100 のストレージコントローラは、有効期間情報を受領し、コピー再開要求を受け付けた時刻から、当該コピーペアの記憶装置 200 側におけるジャーナル有効期限時刻をコピーペアごとに算出する(ステップ 033)。例えば、受領した有効期間情報が示す当該コピーペアのコピー先記憶領域におけるジャーナル有効期間が 10 時間であり、コピー回復要求の受付時刻が 20:00 であった場合、副側ジャーナルの有効期限時刻は 10:00 である。なおジャーナル有効期限時刻はコピーペアごとに設定されている。

【0060】

ストレージコントローラ 160 は、ジャーナル情報管理テーブル 137 に記録されたジャーナルのうち、転送状況が「未転送」であるジャーナルを抽出する。次に、ストレージコントローラ 160 は、この抽出したジャーナル管理情報のうち、各ジャーナル管理情報のデータ記憶領域識別子 1372 で識別されるデータ記憶領域のコピー先記憶領域をペア構成管理テーブルから特定する。そして、ストレージコントローラは、特定したコピー先記憶領域のジャーナル有効期限時刻とジャーナル取得時刻を比較する(ステップ 034)。ステップ 034 の結果、ストレージコントローラは、コピー先記憶領域のジャーナル有効期限時刻よりも前の時点に取得した未転送ジャーナルがある場合(ステップ 034 の Yes)は、有効期限時刻を過ぎた未転送ジャーナルのうち、データ記憶領域 120 内のアドレスごとに、同一アドレスへの書き込みに該当するジャーナルの最新ジャーナルを選抜して記憶装置 200 に転送する(ステップ 035)。ジャーナル転送はステップ 013 から 014 c に前述した処理と同様の手順であればよい。

30

40

【0061】

一方、ステップ 034 の結果、有効期限時刻以降に取得したジャーナルについては、ストレージコントローラ 160 は、書き込み順序番号の順に転送する(ステップ 036)。ストレージコントローラ 160 は、コピーを完了すると、ジャーナル情報管理テーブル 137 における、ステップ 035 とステップ 036 で転送したジャーナルの「転送状況」を「転送済み」に更新する(ステップ 037)。

【0062】

なお、ステップ 035 で同一アドレスへの書き込みに該当するジャーナルのうち、最新のジャーナルを選抜して転送した場合には、ストレージコントローラ 160 は、最新以外

50

の同一アドレスへの書き込みに該当するジャーナルの「転送状況」を、「未転送」の状態から「転送済み」または「転送不要」と変更し、ジャーナル情報管理テーブルを更新する。最後に、ストレージコントローラ160は、コピー状態管理テーブル135における当該コピーペアの「コピー状態」を「正常」に更新する(ステップ038)。以上で、コピー再開処理を完了し、図11におけるステップ001の状態に遷移する。

#### 【0063】

図15は、記憶装置100または記憶装置200のストレージコントローラ160が実行するデータ復元プログラム134でのデータ復元処理の流れを示す。管理者は入力装置を操作し、復元対象のデータ記憶領域と復元時刻を入力する。管理計算機300は、復元対象のデータ記憶領域と復元時刻の入力を受けると、管理計算機300は、復元対象のデータ記憶領域と復元時刻を含むデータ復元要求メッセージを記憶装置100または記憶装置200に送信する(ステップ040a)データ復元要求メッセージは、少なくとも、データ復元対象の記憶領域を示す記憶領域の識別子と、データ復旧指定時刻とを含む記載しておく。

10

#### 【0064】

ストレージコントローラ160は、データ復旧要求を受け、データ復元プログラム138を読み出し、実行を開始する(ステップ041)。データ復元プログラム138を実行するストレージコントローラ160は、スナップショット管理テーブル139を参照し、復元対象のデータ記憶領域にデータ記憶領域識別子1392が一致するスナップショットを抽出する。さらに、ストレージコントローラ160は、これらのスナップショットのうち、指定復元時刻に一致するか、指定時刻以前のスナップショットのうち最新の時点に取得されたスナップショットを特定する(ステップ042)。

20

#### 【0065】

ステップ042の結果、指定復元時刻とスナップショット取得時刻が一致する場合(ステップ043のYes)、ストレージコントローラ160は、そのスナップショットの記憶領域識別子1391を管理情報入出力プログラム131に伝達する(ステップ045)。

#### 【0066】

一方、指定復元時刻とスナップショット取得時刻が一致しない場合(ステップ043のNo)、ストレージコントローラ160は、指定復元時刻とスナップショット取得時刻を比較し、指定復元時刻よりスナップショット取得時刻が後であるか、判定する(ステップ0435)。判定結果、指定復元時刻よりスナップショット取得時刻が後である場合は(ステップ0435のYes)、ストレージコントローラ160は、そのスナップショットの記憶領域識別子1391を管理情報入出力プログラム131に伝達する(ステップ045)。

30

#### 【0067】

一方、ステップ0435の判定結果がNoの場合は、ストレージコントローラは、スナップショット取得時刻からデータ復旧指定時刻との間の時系列において、スナップショット取得時刻に最も近い時刻に作成されたジャーナルをスナップショットへ反映する(044)。この時、反映したジャーナルが有するデータ書き込み時刻が、更新後の当該スナップショットの取得時刻となる。

40

#### 【0068】

ジャーナルを反映後のスナップショット取得時刻が指定復元時刻と一致するか、スナップショット取得時刻が指定復元時刻より後になった場合(043、あるいは0435のYes)、ストレージコントローラ160は、管理情報入出力プログラム131を実行し、ジャーナル反映後のスナップショットのアドレス情報を用いて、データを復元する(045)。以上のように、ストレージコントローラ160は、データ復旧指定時刻とスナップショット取得時刻が一致あるいは以降になるまで一連の操作を繰り返し、データを復元する。

#### 【0069】

50

なお、本実施形態では指定復元時刻より古いスナップショットに新しいジャーナルデータを順に反映させる実行手順を例示したが、逆にステップ042で指定復元時刻以降で最も古いスナップショットを取得し、スナップショット取得時刻が指定復元時刻と一致するかその前になるまで、同スナップショット取得時刻以前のジャーナルを書き込み順序番号とは逆順で反映させていく手順としてもよい。

【0070】

また、ストレージコントローラは、データ復旧指定時刻に最も近い時点に取得されたスナップショットの複製をデータ記憶領域120又はその他の記憶領域に作成し、本複製スナップショットに対して一連のデータ復旧処理を行ってもよい。

【0071】

さらに、管理計算機300において、データ記憶領域120に関する、ジャーナル有効期間情報や、データ復元可能時点の範囲を、管理者に通知するために出力装置330に出力する例を示す。

【0072】

図16は、管理計算機300における出力画面の例を示す。

【0073】

管理計算機300のCPU360は、コピーペアの指定を受けると、ジャーナル有効期間管理テーブル322を参照し、該当するコピーペア識別子3220に対応付けられる正サイトにおけるジャーナル有効期間3221と、副サイトにジャーナル有効期間3222を取得し、それぞれ、表示エリア1600に出力する。次に、CPU360は、記憶装置100及び200からジャーナル情報管理テーブルを定期的を取得し、指定されたコピーペアに該当するデータ記憶領域に対応付けられる取得したジャーナル情報管理テーブルのジャーナル取得時刻から各サイトでの復元可能時間を算出する。

【0074】

CPU360は、算出した結果を保持し、正サイト側のコピー元記憶領域と、副サイト側のコピー先記憶領域のそれぞれについて、ジャーナルによるデータ復元可能時間の実績を時系列で、表示エリア1640に出力する。

【0075】

出力した例が図16に示すとおりで、あるコピーペアを構成する、正サイト側のコピー元記憶領域と、副サイト側のコピー先記憶領域のそれぞれについて、ジャーナル有効期間の設定値と、ジャーナルによるデータ復元可能時間の実績を時系列で表示エリア1640に出力している。

【0076】

また、管理計算機300は、各サイトでの復元可能時間を定期的に監視し、算出した復元可能時間の値が、一定期間、有効期間を下回った場合、表示エリア1620のようなメッセージを出力してよい。また、そのメッセージとあわせて、表示エリア1640の該当する部分を1660や1680のような強調表示してもよい。

【0077】

このように稼働実績を出力することにより、(1)ジャーナル有効期間の設定値に対して、実際のジャーナルが同値を達成するように蓄積されているかどうか、(2)コピー停止を原因としてコピー先記憶領域のデータ復元可能期間が短縮されていないかどうかといった分析判断が可能となる。管理計算機300は、ジャーナルの状況を管理者に向けて出力することで、その今後のシステム構築や設計変更を検討するなどの運用管理を手助けする。

【0078】

以上により、記憶装置間でコピー処理が停止した場合においても、記憶装置や記憶領域それぞれのジャーナルの保持期間に応じたコピー制御が可能となる。

【0079】

以下に、本実施形態における運用例を解説する。

【0080】

10

20

30

40

50

図5のペア構成管理テーブル135において、コピーペア識別子01であるコピーペアは、記憶装置識別子1000の記憶装置内のデータ記憶領域00:01をコピー元データ記憶領域とし、記憶装置識別子2000の記憶装置内のデータ記憶領域0A:01をコピー先データ記憶領域とする。

【0081】

図6のコピー状態管理テーブル136に示すように、コピーペア識別子01のコピーペアに対するデータコピー処理が、データ転送用ネットワークの障害によって時刻09:30に「停止」に更新されたと想定する。

【0082】

コピーペア01のデータコピー処理が停止した後も、データ記憶領域00:01への書き込み処理は継続され、書き込み処理に対し、ジャーナル取得プログラム132によってジャーナル(ジャーナルデータ及びジャーナルメタデータ)が取得・蓄積される。

10

【0083】

管理者は、障害要因を取り除いた後、管理計算機300は、時刻20:00にコピーを再開するコピー再開要求を記憶装置100に送信する。なお、コピーを再開する設定や再開する時刻の指定は、管理者の操作により入力装置340から入力されてもよい。

【0084】

記憶装置100において、コピー再開要求を受領すると、データコピープログラム133がデータコピー再開処理を実施する。データコピー再開処理において転送するジャーナルは、コピー再開要求を発信した時刻20:00から、コピー先のジャーナル有効期間である10時間遡った時刻、すなわち、時刻10:00よりも前に取得したジャーナルの最新のもを、更新されたデータ記憶領域内のアドレス毎に選択し、転送する。また、時刻10:00以降に取得したジャーナルについては、書き込み順序番号の順に記憶装置200に転送する。データコピー再開処理を終了すると、コピー状態は「正常」となり、通常のデータコピー処理(データ記憶領域への書き込み処理ごとに取得したジャーナルを、書き込み処理とは非同期に転送する処理)を継続する。

20

【0085】

次に、例えば、データコピー再開処理完了後、管理者が、コピーペア01の時刻19:00時点のデータを復元する要求を管理計算機から記憶装置200に対して送信した場合を説明する。記憶装置200内のデータ復元プログラム134は、スナップショット管理テーブル139を参照し、データ復元指定時刻である19:00以前であってこれに最も近い時点において取得されたスナップショットを特定する。本例では、図8に示したスナップショット時刻情報18:00を有するスナップショットがこれに該当する。データ復元プログラム134は、本スナップショットに、時刻18:00以降に取得されたジャーナルを、ジャーナルデータ記憶領域121及びジャーナルメタデータ記憶領域122から読み出し、ジャーナルメタデータに含まれる情報を参照して当該スナップショットに書き込む。具体的には、図7に示したジャーナル情報管理テーブルの書き込み番号05で表されるジャーナルを書き込むことにより、時刻19:00におけるデータを復元する。

30

【0086】

以上が、本実施形態における運用例である。

40

【0087】

次に、実施例1の変形例について図17、18、19、20を用いて説明する。実施例1と異なる点は、図14に示したデータコピー再開時の処理が、以下に解説する図20のデータコピー再開処理に置き換わる点である。変形例における記憶システムの構成機器及び接続形態は、図1に示した第一の実施例のものと同じでよい。また、記憶装置100及び記憶装置200の構成についても、図2に示した第一の実施例のものと同じでよい。

【0088】

図17は、管理計算機300のプログラムメモリ320が保持するプログラム群及びテーブル群を示す。管理情報入出力プログラム321及びジャーナル有効期間管理テーブル322は、実施例1に示したものと同一でよい。

50

## 【 0 0 8 9 】

転送ジャーナル管理テーブル 3 2 4 は、記憶装置 1 0 0 においてデータコピー処理を再開する際に、記憶装置 2 0 0 に転送するジャーナルの情報を蓄積する。データコピープログラム 1 3 3 を実行する記憶装置 1 0 0 のストレージコントローラ 1 6 0 は、データコピー処理の再開時、本テーブルに記載されたジャーナルを転送する。ペア構成管理テーブル 1 3 5 は、図 5 に記載したものと同様であればよい。本実施形態においては、随時記憶装置 1 0 0 と通信し、ペア構成管理テーブル 1 3 5 を更新するために、記憶装置 1 0 0 からペア構成やペア状態を問合せする。

## 【 0 0 9 0 】

図 1 8 は、転送ジャーナル管理テーブル 3 2 4 の例を示す。図 1 8 において、転送ジャーナル管理テーブル 3 2 4 は、少なくともジャーナルの書き込み順序番号 3 2 4 0、データ記憶領域識別子 3 2 4 1、データ記憶領域内アドレス情報 3 2 4 2、ジャーナルデータ記憶領域識別子 3 2 4 3、ジャーナルメタデータ記憶領域識別子 3 2 4 4 を有する。書き込み順序番号 3 2 4 0、データ記憶領域識別子 3 2 4 1、データ記憶領域内アドレス情報 3 2 4 2、ジャーナルデータ記憶領域識別子は、図 7 に前述したジャーナル情報管理テーブル 1 3 7 と同様のパラメータである。さらにジャーナルメタデータ記憶領域識別子 3 2 4 4 は、ジャーナル情報管理テーブル 1 3 7 を格納する記憶領域を特定する識別子である。

10

## 【 0 0 9 1 】

ジャーナル監視プログラム 3 2 2 は、記憶装置 2 0 0 内のジャーナル記憶領域に格納されたジャーナルの有効期間について管理するプログラムである。具体的には、記憶装置 2 0 0 内のジャーナル情報管理テーブルの情報を定期的に取得し、コピー回復処理時に記憶装置 1 0 0 から記憶装置 2 0 0 に転送するジャーナルに関する情報を管理する。

20

## 【 0 0 9 2 】

図 1 9 は、ジャーナル監視プログラム 3 2 2 を実行する管理計算機 3 0 0 の CPU 3 6 0 の処理を示す。まず、CPU 3 6 0 は、記憶装置 1 6 0 にジャーナル情報管理テーブルを要求する（ステップ 1 4 0 0）。要求を受領した記憶装置 1 6 0 は、ジャーナル情報管理テーブルを管理計算機 3 0 0 に送信する（ステップ 0 3 2）。

## 【 0 0 9 3 】

管理計算機 3 0 0 の CPU 3 6 0 は、ジャーナル情報管理テーブルを受領する。次に、CPU 3 6 0 は、ペア構成管理テーブルを参照し、ジャーナル情報管理テーブルのデータ記憶領域識別子 3 2 4 1 に関連付けられているコピーペア識別子 1 3 5 0 及びコピー先記憶装置識別子 1 3 5 3 を特定する。そして、CPU 3 6 0 は、特定したコピー識別子に対応する各コピーペアについて、ジャーナル有効期間管理テーブル 3 2 2 を参照し、該当するジャーナル有効期間と現在時刻とから副サイトにおけるジャーナル期限時刻を算出する（ステップ 1 4 0 1）。例えば、現在時刻が 1 0 : 0 0 であり、ジャーナル有効期間が 1 0 時間である場合、ジャーナル有効期限時刻は、1 0 時間前の 0 : 0 0 となる。なおジャーナル有効期限時刻はコピーペアごとに設定されている。

30

## 【 0 0 9 4 】

そして、プロセッサ 3 6 0 は、記憶装置 1 0 0 から取得したジャーナル情報管理テーブル 1 3 7 のうち、転送状況 1 3 7 5 が「未転送」のジャーナル管理情報を抽出する（ステップ 1 4 0 6）。次に、プロセッサ 3 6 0 は、ステップ 1 4 0 6 で抽出したジャーナル管理情報のうち、各ジャーナル管理情報のデータ記憶領域識別子 3 2 4 1 で識別されるデータ記憶領域のコピー先記憶領域をペア構成管理テーブルから特定し、同コピー先記憶領域のジャーナル有効期限時刻よりも前の時点に取得したジャーナルのみ抽出する（ステップ 1 4 0 2）。さらに、プロセッサ 3 6 0 は、ステップ 1 4 0 2 で抽出したジャーナル管理情報のうち、それぞれのデータ記憶領域について、同データ記憶領域内の同一アドレスへの書き込みに該当するジャーナルがあれば、最新のジャーナルを残し、それ以外のジャーナル管理情報を取り除いた抽出結果を作成する（ステップ 1 4 0 3）。以上のステップにより、有効期限時刻より前のジャーナルについては、未転送かつ最新の情報に絞り込んだ

40

50



抽出結果を取得したこととなる。

【0095】

さらに、CPU360は、記憶装置100から取得したジャーナル管理情報から、ジャーナル有効期限時刻以後に取得したジャーナル情報を抽出する(ステップ1404)。CPU360は、ステップ1403および1404で特定したジャーナル情報に基づき、図18に示す転送ジャーナル管理テーブル324を作成する(ステップ1405)。

【0096】

なお、ステップ1400を行なうタイミングは、コピーに関する障害を取り除いた後や記憶装置200の計画停止やコピー回線のメンテナンスが完了した後に行なってもよい。

図20は、本変形例におけるデータコピー再開時の処理の流れを示す。まず、管理者が管理計算機300の入力装置を介してコピー再開要求操作を入力する。管理計算機300のCPU360は、入力を受けたコピー再開要求メッセージを、転送ジャーナル情報管理テーブル324とあわせて記憶装置100に送信する(ステップ1500)。

【0097】

コピープログラム133を実行するストレージコントローラ160は、転送ジャーナル情報管理テーブル324に記録されたジャーナルそれぞれについて、ジャーナルに付与された書き込み順序番号の順に従って、記憶装置200に転送する(ステップ1503)。

【0098】

ストレージコントローラ160は、記憶装置200からコピー完了通知を受領すると、ジャーナル情報管理テーブル137における、転送したジャーナルの「転送状況」を「転送済み」に更新する(ステップ1504)。このとき、転送したジャーナルのデータ記憶領域内の同一アドレスへの古い時点の書き込みがあり、転送状況が「未転送」であった場合は、同ジャーナル管理情報の転送状況を「転送済み」あるいは「転送不要」と更新しておく。

【0099】

次に、ストレージコントローラ160は、コピー状態管理テーブル136における当該コピーペアの「コピー状態」を「正常」に更新する(ステップ1505)し、これを以って、コピー再開処理が終了する。

【0100】

なお、変形例で管理計算機300から記憶装置100に図18の代わりに、図7のジャーナル情報管理テーブルを更新したものを記憶装置100に送信してもよい。すなわち、ステップ1403で、抽出結果を作成せずに、図7のジャーナル情報管理テーブル137の「転送状況」1375にそれ以外のジャーナル管理情報に、転送不要であることを示すフラグを対応付け、CPU360は、フラグを対応付けたジャーナル管理情報を記憶装置100に送信してもよい。

【0101】

変形例により、管理計算機が、記憶装置100及び記憶装置200のジャーナル保持期間に応じたコピー制御を指示することができる。または、管理計算機が、コピーを再開するにあたり、記憶装置200の属性に応じたコピー制御を指示することができる。

【図面の簡単な説明】

【0102】

【図1】本発明の第一の実施形態における計算機システムの構成を示した図である。

【図2】本発明の第一の実施形態における記憶装置内の構成を模式的に示した図である。

【図3】本実施形態におけるデータコピー再開時のジャーナル転送範囲を模式的に示した図である。

【図4】本実施形態における任意時点のデータ復元方法を模式的に示した図である。

【図5】記憶装置が格納する、ペア構成管理テーブルの例である。

【図6】記憶装置が格納する、コピー状態管理テーブルの例である。

【図7】記憶装置が格納するジャーナル情報管理テーブルの例である。

【図8】記憶装置が格納する、スナップショット管理テーブルの例である。

10

20

30

40

50

【図 9】本発明の第一の実施形態における管理計算機内の構成を模式的に示した図である。

【図 10】管理計算機が格納するジャーナル有効期間管理テーブルの例である。

【図 11】本実施形態における計算機システムの運用の流れを表すフロー状態遷移図である。

【図 12】本実施形態におけるコピー停止処理フローを示した図である。

【図 13】本実施形態におけるコピー状態が「停止」である場合のジャーナル管理のフローを示す。

【図 14】本実施形態におけるコピー再開時の処理フローを示した図である。

【図 15】本実施形態におけるデータ復元処理のフローを示した図である。

【図 16】本実施形態における出力画面の例である。

【図 17】変形例における管理計算機内の構成を模式的に示した図である。

【図 18】変形例における管理計算機が格納する、転送ジャーナル管理テーブルの例である。

【図 19】変形例における管理計算機が実行する、ジャーナル監視プログラムの処理フローを示した図である。

【図 20】変形例における、コピー再開時の処理フローを示した図である。

【図 21】本実施形態における運用管理操作によるコピー停止処理フローを表す。

【符号の説明】

【0103】

100 記憶装置

120 データ記憶領域

121 ジャーナルデータ記憶領域

122 ジャーナルメタデータ記憶領域

123 スナップショット記憶領域

131 管理情報入出力プログラム

132 ジャーナル取得プログラム

133 データコピープログラム

134 データ復元プログラム

135 ペア構成情報管理テーブル

136 コピー状態管理テーブル

137 ジャーナル情報管理テーブル

160 ストレージコントローラ

200 記憶装置

300 管理計算機

322 ジャーナル有効期間管理テーブル

330 出力装置

340 入力装置

350 管理用通信インタフェース

360 CPU

400 ホスト計算機

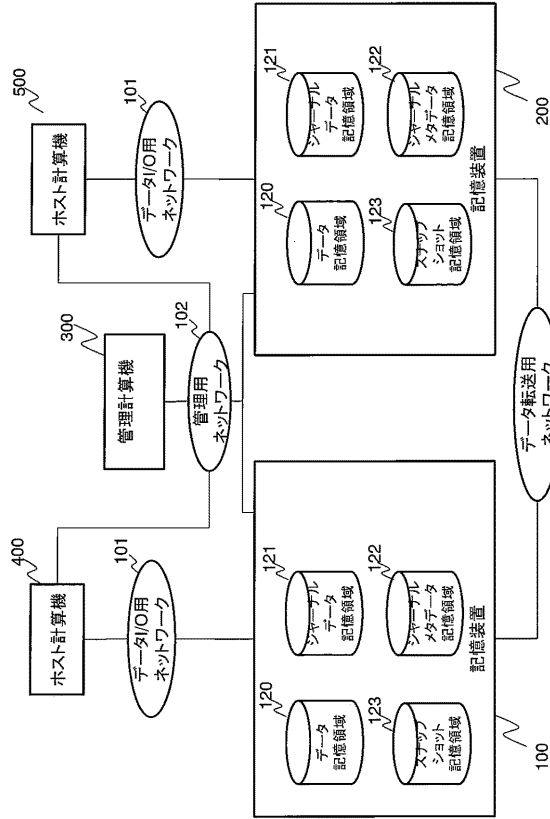
10

20

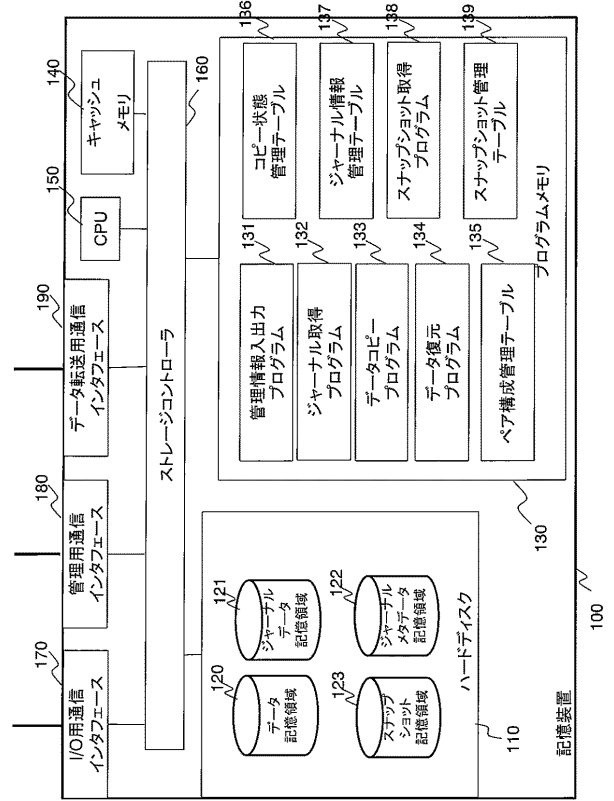
30

40

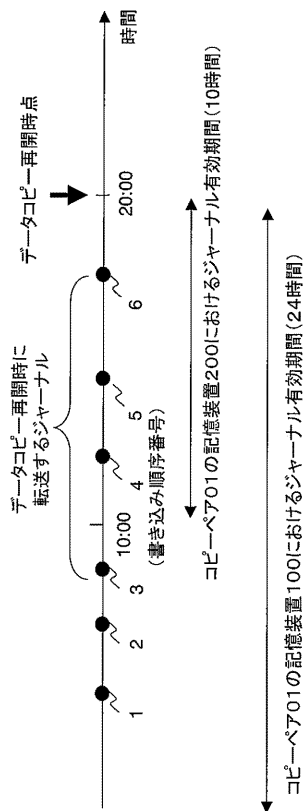
【 図 1 】



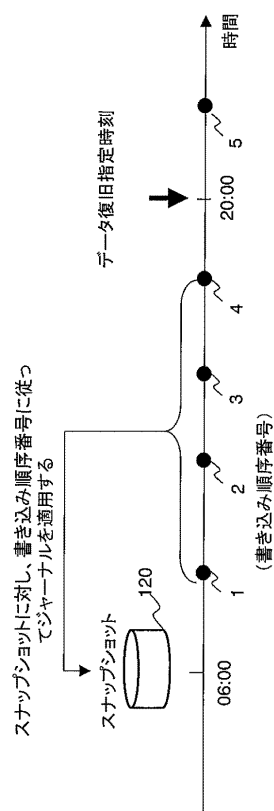
【 図 2 】



【 図 3 】



【 図 4 】



【 図 5 】

図5  
ペア構成管理テーブル

コピーペア 識別子	コピー元 記憶装置識別子	コピー元データ 記憶領域識別子	コピー先 記憶装置識別子	コピー先データ 記憶領域識別子
01	1000	00:01	2000	00A:01
02	1100	00:02	2100	00A:02
03	1200	00:03	2200	00A:03

【 図 6 】

図6  
コピー状態管理テーブル

コピーペア 識別子	コピー状態	コピー状態更新時刻
01	停止	09:30:00
02	正常	12:00:00
03	停止	12:05:00

【 図 7 】

図7  
ジャーナル情報管理テーブル

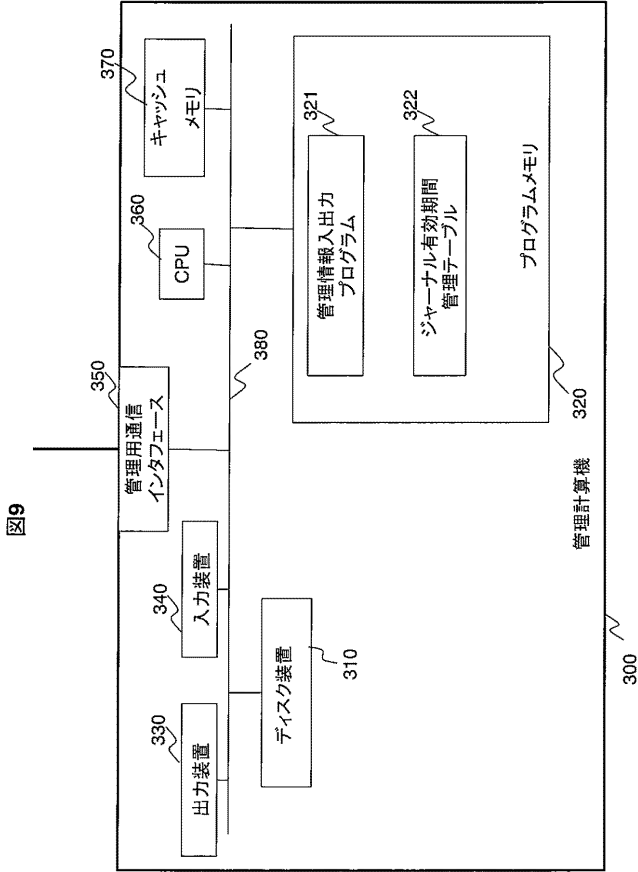
書き込み順 番号	ジャーナル取得時刻	データ記憶領域 識別子	データ記憶領域 内のアドレス	ジャーナルデータ 記憶領域識別子	転送 状況
01	08:08:00	00:01	F5	0B:01	転送済み
02	09:40:00	00:02	E0	0B:02	未転送
03	09:50:00	00:01	A8	0B:01	未転送
04	10:15:00	00:01	D7	0B:01	未転送
05	18:05:00	00:02	A1	0B:02	未転送
06	20:00:00	00:01	D2	0B:01	未転送

【 図 8 】

図8  
スナップショット管理テーブル

スナップショット 時刻情報	スナップショット 記憶領域識別子	データ 記憶領域識別子	ジャーナルデータ 記憶領域識別子	ジャーナルメタデータ 記憶領域識別子
09:00	10:01	00:01	0B:01	0C:01
14:00	10:02	00:02	0B:02	0C:02
18:00	10:03	00:03	0B:03	0C:03

【 図 9 】



【 図 1 1 】

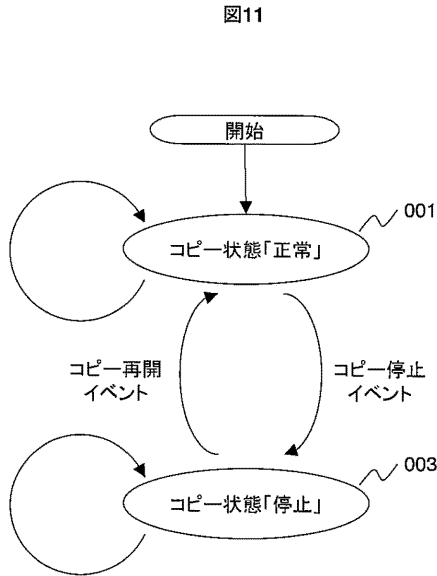


図11

【 図 1 0 】

図10  
ジャーナル有効期間管理テーブル

コピー識別子	正サイトにおけるジャーナル有効期間	副サイトにおけるジャーナル有効期間
01	24時間	10時間
02	168時間	24時間
03	168時間	10時間

【 図 1 2 】

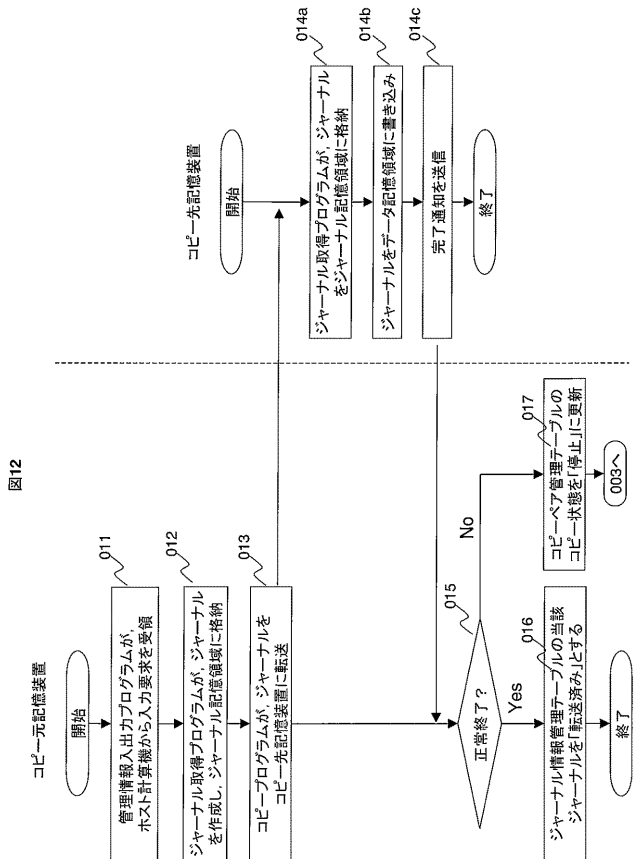
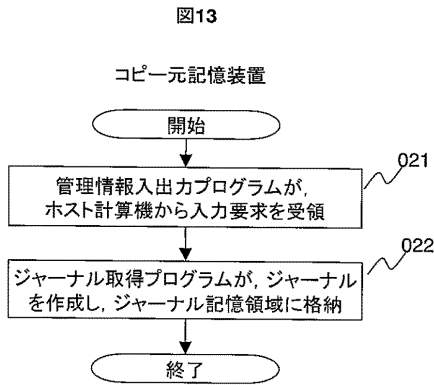
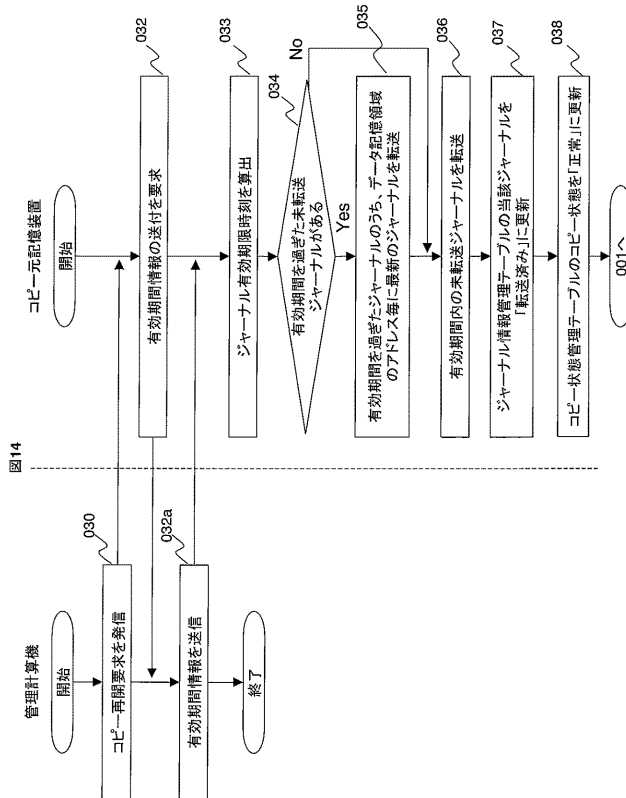


図12

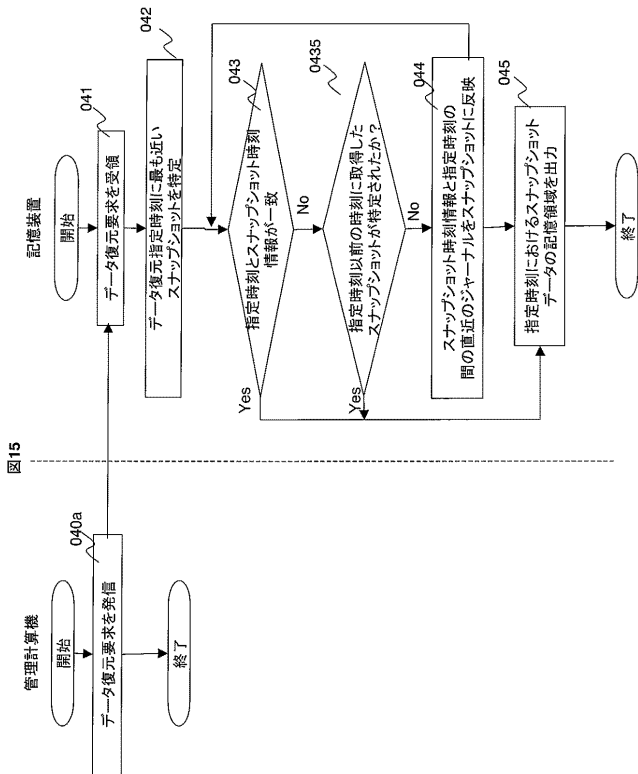
【 図 1 3 】



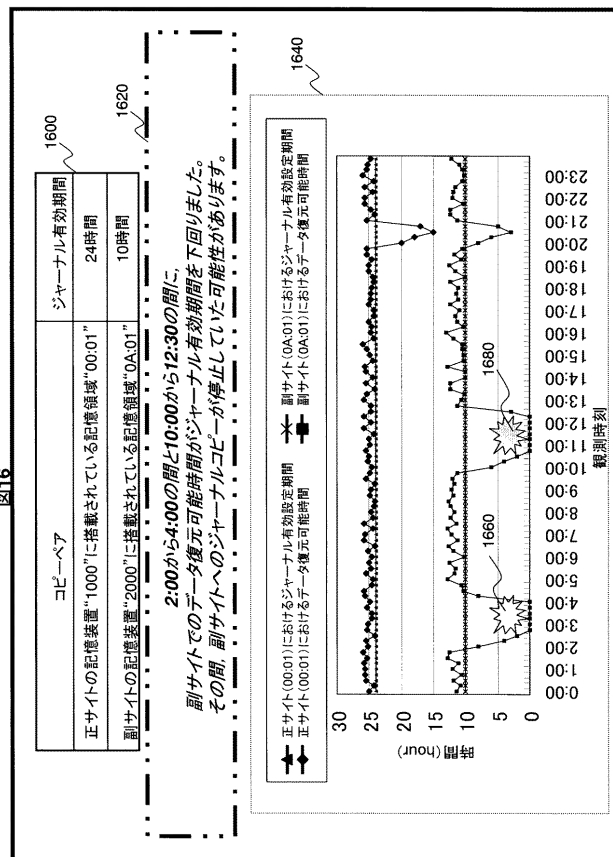
【 図 1 4 】



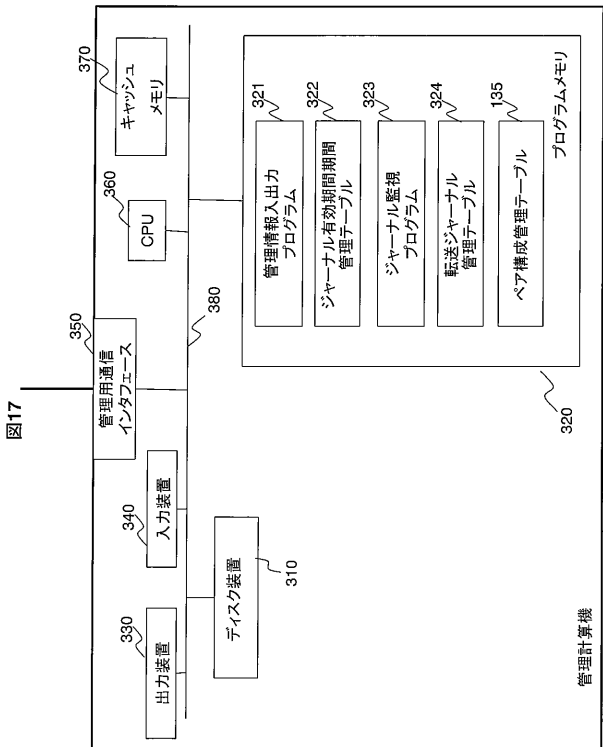
【 図 1 5 】



【 図 1 6 】



【 図 1 7 】

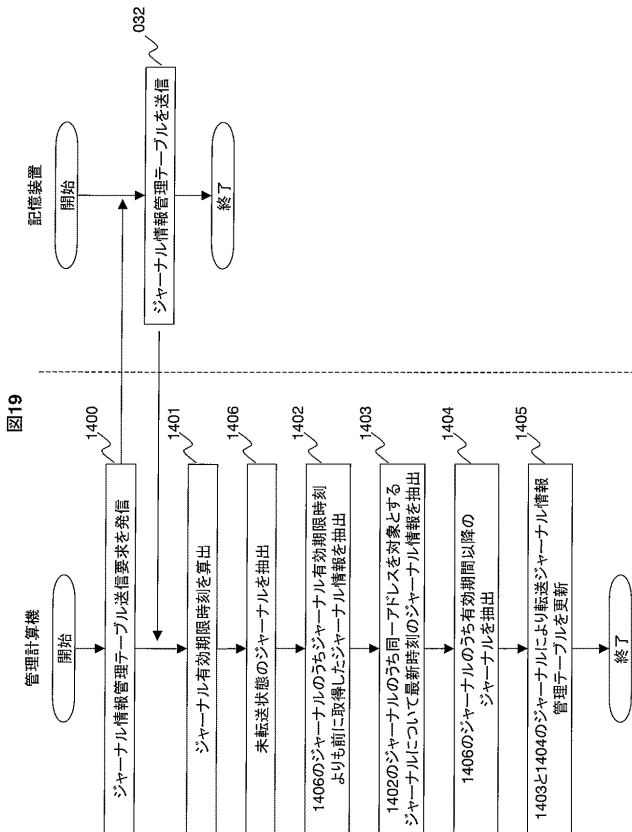


【 図 1 8 】

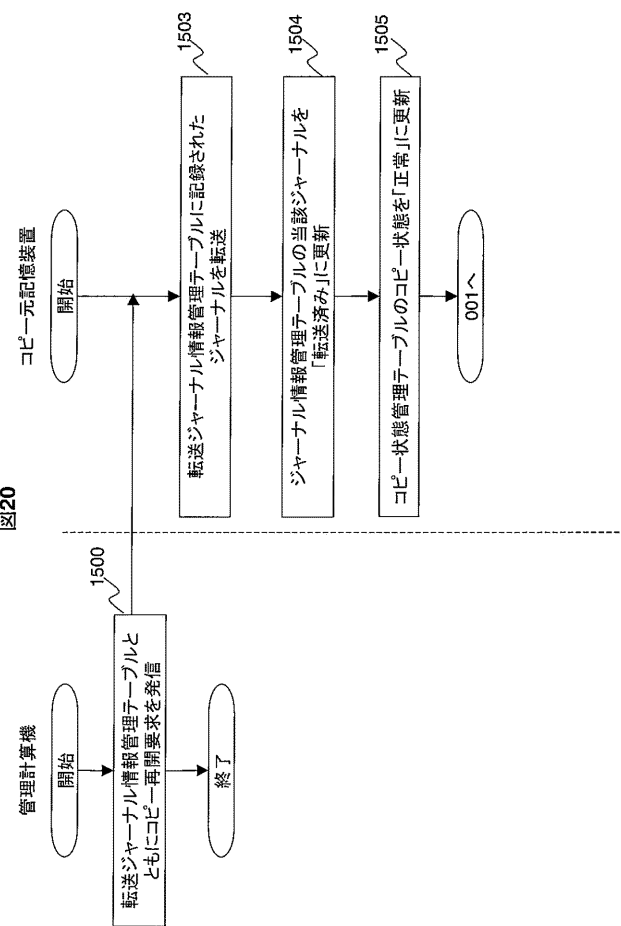
【図18】 転送ジャーナル管理テーブル

書き込み順序番号	データ記憶領域識別子	データ記憶領域アドレス情報	ジャーナルデータ記憶領域識別子	ジャーナルメタデータ記憶領域識別子
03	00:01	A8	0B:01	0C:01
04	00:01	D7	0B:01	0C:01
05	00:03	A1	0B:02	0C:02
06	00:01	D2	0B:01	0C:01

【 図 1 9 】

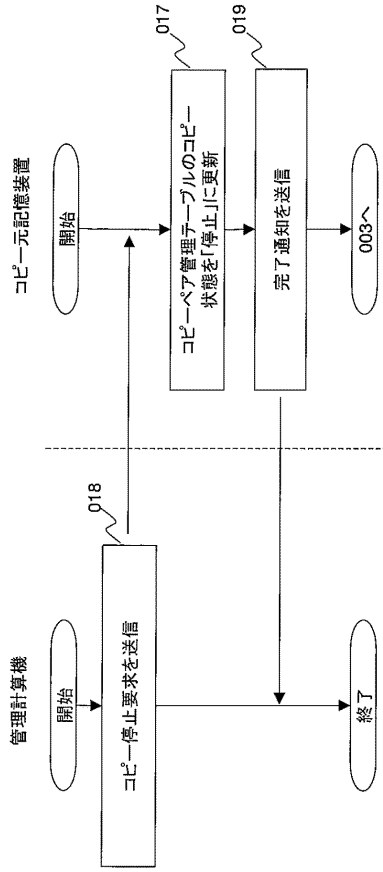


【 図 2 0 】



【 図 2 1 】

図21





---

フロントページの続き

- (72)発明者 田口雄一  
神奈川県川崎市麻生区王禅寺 1 0 9 9 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 山本政行  
神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内
- Fターム(参考) 5B065 BA01 CE02 CE21  
5B082 DD02 DE04