

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6344798号
(P6344798)

(45) 発行日 平成30年6月20日 (2018. 6. 20)

(24) 登録日 平成30年6月1日 (2018. 6. 1)

(51) Int. Cl.	F I		
G06F 3/06 (2006.01)	G06F	3/06	304F
G06F 12/00 (2006.01)	G06F	3/06	302A
	G06F	12/00	514K
	G06F	12/00	531M

請求項の数 11 (全 38 頁)

(21) 出願番号	特願2015-527787 (P2015-527787)	(73) 特許権者	504277388
(86) (22) 出願日	平成25年11月15日 (2013. 11. 15)		▲ホア▼▲ウェイ▼技術有限公司
(65) 公表番号	特表2015-527670 (P2015-527670A)		HUAWEI TECHNOLOGIES
(43) 公表日	平成27年9月17日 (2015. 9. 17)		CO., LTD.
(86) 国際出願番号	PCT/CN2013/087229		中華人民共和国518129広東省深▲セ
(87) 国際公開番号	W02015/010394		ン▼市龍岡区坂田華為本社ビル
(87) 国際公開日	平成27年1月29日 (2015. 1. 29)		Huawei Administrati
審査請求日	平成26年10月22日 (2014. 10. 22)		on Building, Bantian
審判番号	不服2017-712 (P2017-712/J1)	(74) 代理人	100146835
審判請求日	平成29年1月18日 (2017. 1. 18)		弁理士 佐伯 義文
(31) 優先権主張番号	PCT/CN2013/080203	(74) 代理人	100140534
(32) 優先日	平成25年7月26日 (2013. 7. 26)		弁理士 木内 敬二
(33) 優先権主張国	中国 (CN)		

最終頁に続く

(54) 【発明の名称】 データ送信方法、データ受信方法、及びストレージデバイス

(57) 【特許請求の範囲】

【請求項1】

少なくとも2つの他のストレージデバイスの1つに複製が行われる現在の複製タスクがトリガされるとき、カレント時間間隔番号を第1のストレージデバイスによって変更するステップと、

前記第1のストレージデバイスによって、前記変更されたカレント時間間隔番号を読み出すステップと、

前記第1のストレージデバイスによって、第2の番号を読み出すステップであって、前記第2の番号は、前記現在の複製タスクと同じ複製関係に属する最新の完了した複製タスクに対応する番号である、ステップと、

前記第1のストレージデバイスによって、前記変更されたカレント時間間隔番号及び前記第2の番号に従って、第1の番号を決定するステップであって、前記第1の番号は、前記現在の複製タスクがトリガされるときに読み出された前記変更されたカレント時間間隔番号より前の番号であるとともに、前記第1の番号は、前記第2の番号より後の番号である、ステップと、

前記第1のストレージデバイスによって、複製されるべきデータ及び前記複製されるべきデータのアドレス情報を第2のストレージデバイスに複製するステップであって、前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報は前記第1のストレージデバイスのキャッシュに格納されている、かつ前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報は前記第1の番号に対応する、ステップと

を有することを特徴とするデータ複製方法。

【請求項 2】

前記現在の複製タスクがトリガされる時、前記方法は、前記複製関係に対応する識別子を受信するステップをさらに有し、

第2の番号を読み出す前記ステップは、前記識別子に従って前記第2の番号を読み出すステップを有することを特徴とする請求項 1 に記載の方法。

【請求項 3】

現在の複製タスクがトリガされる前に、前記方法は、

第1のデータ書込み要求を受信するステップであって、前記第1のデータ書込み要求は、前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報を有する、ステップと、

前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報に、前記第1の番号を付与するとともに、前記第1の番号が付与された前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報を前記キャッシュに書き込むステップであって、前記第1の番号は、履歴時間期間番号である、ステップと

をさらに有することを特徴とする請求項 1 又は 2 に記載の方法。

【請求項 4】

前記カレント時間間隔番号は、前記履歴時間期間番号を変更することによって形成されることを特徴とする請求項 3 に記載の方法。

【請求項 5】

前記第1のストレージデバイスのキャッシュに格納されている、かつ第1の番号に対応する、複製されるべきデータ及び前記複製されるべきデータのアドレス情報を第2のストレージデバイスに複製する前記ステップは、

前記アドレス情報が複数の番号に対応するとき、前記アドレス情報に対応する前記番号のうちの最新の番号が第1の番号であることを決定するステップと、

前記第1のストレージデバイスのキャッシュに格納されている、かつ前記第1の番号に対応する、前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報を前記第2のストレージデバイスに複製するステップと

を有することを特徴とする請求項 1 から 4 のいずれか一項に記載の方法。

【請求項 6】

ストレージデバイスであって、

少なくとも2つの他のストレージデバイスの1つに複製が行われる現在の複製タスクがトリガされる時、変更されたカレント時間間隔番号を読み出すとともに、第2の番号を読み出すように構成される読み出し及び書込みモジュールであって、前記第2の番号は、前記現在の複製タスクと同じ複製関係に属する最新の完了した複製タスクに対応する番号である、読み出し及び書込みモジュールと、

前記読み出されたカレント時間間隔番号及び前記第2の番号に従って、少なくとも1つの第1の番号を決定するように構成される決定モジュールであって、前記第1の番号は、前記現在の複製タスクがトリガされる時に読み出されたカレント時間間隔番号より前の番号であるとともに、前記第1の番号は、前記第2の番号より後の番号である、決定モジュールと、

複製されるべきデータ及び前記複製されるべきデータのアドレス情報を第2のストレージデバイスに複製するように構成される複製モジュールであって、前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報は前記ストレージデバイスのキャッシュに格納されている、かつ前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報は前記第1の番号に対応する、複製モジュールと

を具備することを特徴とするストレージデバイス。

【請求項 7】

前記ストレージデバイスは、受信モジュールをさらに具備し、

前記受信モジュールは、前記複製関係に対応する識別子を受信するように構成され、

10

20

30

40

50

前記読出し及び書込みモジュールは、前記識別子に従って前記第2の番号を読み出すように特に構成されることを特徴とする請求項6に記載のストレージデバイス。

【請求項8】

前記受信モジュールは、現在の複製タスクがトリガされる前に、第1のデータ書込み要求を受信するようにさらに構成され、

前記第1のデータ書込み要求は、前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報を有し、

前記読出し及び書込みモジュールは、前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報に、前記第1の番号を付与するとともに、前記第1の番号が付与された前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報を前記キャッシュに書き込むようにさらに構成され、前記第1の番号は、履歴時間期間番号であることを特徴とする請求項7に記載のストレージデバイス。

10

【請求項9】

前記カレント時間間隔番号は、前記履歴時間期間番号を変更することによって形成されることを特徴とする請求項8に記載のストレージデバイス。

【請求項10】

前記複製モジュールは、前記アドレス情報が複数の番号に対応するとき、前記アドレス情報に対応する前記番号のうちの最新の番号を第1の番号として決定するとともに、前記キャッシュに格納されている、かつ前記第1の番号に対応する、前記複製されるべきデータ及び前記複製されるべきデータのアドレス情報を前記第2のストレージデバイスに複製するように特に構成されることを特徴とする請求項6から9のいずれか一項に記載のストレージデバイス。

20

【請求項11】

請求項1から5のいずれか一項に記載の方法をコンピュータに実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ技術に関し、特にデータ送信方法、データ受信方法、及びストレージデバイスに関する。

【0002】

30

本願は、参照により全体を本明細書に組み込まれる「Data Sending Method, Data Receiving Method, and Storage Device」と題する、2013年7月26日に提出された国際出願PCT/CN2013/080203の継続出願である。

【背景技術】

【0003】

遠隔データ複製技術(remote data replication technologies)としても知られる、データ障害復旧(Data disaster recovery)は、ローカルデータの有効な複製である非ローカルシステムの機構に言及する。ローカルデータ又は全体のアプリケーションシステムにおいて障害が発生するとき、システムの不可欠のサービスデータの少なくとも1つの有効なコピーが非ローカルに格納される。

40

【0004】

一般的なデータ障害復旧システムは、生産センタ及び障害復旧センタを含む。生産センタでは、ホスト及びストレージレイがサービスの正常動作のために配置され、障害復旧センタでは、ホスト及びストレージレイが、生産センタが障害に遭遇した後に生産センタのサービスを引き継ぐために配置される。生産センタ又は障害復旧センタのいずれかのストレージレイは、複数のデータボリュームを含み、データボリュームは、物理ストレージ領域をマッピングすることによって形成される論理的ストレージ領域である。生産センタでのサービスによって生成されたデータが生産レイに書き込まれた後、そのデータは障害復旧リンクを使用して障害復旧センタに複製され得るとともに、障害復旧レイに書き込まれ得る。障害復旧センタ内のデータが、障害発生後にサービスの引き継ぎをサポ

50

ートできるように、障害復旧アレイに複製されたデータの一貫性が保証されなければならない。データの一貫性の保証は、実際は依存性ベースのデータ書込み要求であり、依存性は、保証される必要がある。アプリケーションプログラム、オペレーティングシステム、及びデータベースは全て、本質的に自身のサービスを実行するためにデータ書込み要求のこの依存性のロジックに頼っている。例えば、データ書込み要求2は、データ書込み要求1が完了するまで実行されない。順番は固定である。換言すると、データ書込み要求1が成功して完全に返されることが確保されるまで、システムはデータ書込み要求2を配信しない。このように、実行プロセスが失敗のために中断されるとき、サービスは特有の方法に頼ることによってリストアされ得る。そうではない場合、例えば、データが読み出されると、データ書込み要求2によって格納されたデータが読み出されることができ一方、データ書込み要求1によって格納されたデータが読み出されることができないとき、結果的にサービスがリストアされ得ないことがあり得る。

10

【0005】

従来技術では、スナップショット技術が問題の解決に使用されている。

スナップショットは、ある時点(time point)(コピー動作が開始された時点)でのデータのイメージである。スナップショットの目的は、特定の時点でのデータボリュームのための状態ビュー(state view)を作成することである。このビューから、作成時のデータボリュームのデータのみを見ることができ一方、その時点の後のデータボリュームへの変更(新しいデータが書き込まれる)は、スナップショットビューに反映されない。このスナップショットビューを使用することによって、データの複製が実施され得る。生産センタにとって、スナップショットデータは、「静的」である。従って、データのスナップショットが各時点で取得された後、生産センタは障害復旧センタにスナップショットデータを複製できる。このことは、遠隔データ複製を実施するだけでなく、生産センタ内の後続のデータ書込み要求の実行に影響を与えない。障害復旧センタにとって、データの一貫性の要求も満足され得る。例えば、データ書込み要求2のデータの障害復旧センタへの複製が成功する一方、データ書込み要求1の複製が成功しないとき、データ書込み要求2より前のスナップショットデータが、障害復旧センタ内のデータを以前の状態にリストアするために使用され得る。

20

【0006】

データ書込み要求を実行するとき、生産センタがスナップショットを処理するとともに、スナップショットデータのストレージ専用のデータボリュームに生成されたスナップショットデータを格納する必要があるので、生産センタは、障害復旧センタにスナップショットデータを複製するとき、データボリュームに格納されたスナップショットデータをキャッシュに読み出すとともに、障害復旧センタにスナップショットデータを送信する必要がある。しかしながら、スナップショットデータを生成するために使用されるデータが、キャッシュに存在したままであり、適切に利用され得ない場合がある。全ての複製は、データボリュームからスナップショットデータを読み出すことを要求し、そのことは長時間のデータ複製と低効率をもたらす。

30

【発明の概要】**【課題を解決するための手段】**

40

【0007】

本発明の実施形態は、データ送信方法を提供し、データ書込み要求において伝送される情報は、第1のストレージデバイスのキャッシュから第2のストレージデバイスに直接送信されることができ、そのことがデータ複製の効率を上げる。

【0008】

第1態様では、本発明の実施形態は、データ送信方法を提供し、前記方法は、ホストによって送信される第1のデータ書込み要求を、第1のストレージデバイスによって受信するステップであって、前記第1のデータ書込み要求は、書き込むべきデータ及びアドレス情報を伝送する、ステップと、第1の番号を、書き込むべきデータ及びアドレス情報に付与するとともに、第1の番号が付与された前記書き込むべきデータ及び前記アドレス情報を

50

キャッシュに書き込むステップであって、前記第1の番号は、カレント時間間隔番号である、ステップと、第1の番号に対応する前記書き込むべきデータ及び前記アドレス情報を、前記キャッシュから読み出すステップと、後続のデータ書き込み要求において伝送される情報を識別するために、前記カレント時間間隔番号を変更するステップと、前記書き込むべきデータ及び前記アドレス情報を第2のストレージデバイスに送信するステップとを含む。

【0009】

本発明の実施形態の第1の態様の第1の可能な実装では、前記第1の番号は、現在の複製タスクを識別するために使用され、前記方法は、第2の番号を記録するステップをさらに含み、前記第2の番号は、前記現在の複製タスクの前の最新の完了した複製タスクに対応する番号である。

10

【0010】

本発明の実施形態の第1の態様の第1の可能な実装を参照すると、第1の態様の第2の可能な実施では、方法は、前記第2の番号の後であるとともに、前記第1の番号に先行する番号に対応する書き込むべきデータ及びアドレス情報を、前記キャッシュから読み出すステップと、前記第2の番号の後であるとともに、前記第1の番号に先行する前記番号に対応する前記書き込むべきデータ及び前記アドレス情報を前記第2のストレージデバイスに送信するステップとをさらに含む。

【0011】

本発明の実施形態の第1の態様の第3の可能な実装では、方法は、前記カレント時間間隔番号を記録するステップをさらに含み、前記カレント時間間隔番号は、前記第1の番号を生成するために使用される。

20

【0012】

第2の態様において、本発明の実施形態はデータ受信方法を提供し、前記方法は、第1のストレージデバイスによって送信されたアドレス情報を、第2のストレージデバイスによって受信するステップと、前記第1のストレージデバイスが故障していることが決定すると、第1の番号に対応する書き込むべきデータを、前記アドレス情報に従って前記第2のストレージデバイスによって取得するステップであって、前記第1の番号に対応するアドレス情報が前記受信したアドレス情報と同じであるとともに、前記第1の番号がカレント時間間隔番号に先行する番号である、ステップと、前記第1の番号に対応する前記書き込むべきデータ及び前記アドレス情報に第2の番号を付与するとともに、情報をキャッシュに書き込むステップとを含む。

30

【0013】

本発明の実施形態の第2の態様の第1の可能な実装として、方法は、前記カレント時間間隔番号を記録するステップをさらに含み、前記カレント時間間隔番号は、前記第2の番号を生成するために使用される。

【0014】

本発明の実施形態の第2の態様の第2の可能な実装では、方法は、ホストによって送信されたデータ読み出し要求を受信するステップであって、前記データ読み出し要求は、前記受信したアドレス情報を有する、ステップと、前記受信したアドレス情報に対応する最新の番号が前記第2の番号であることを決定するステップと、前記ホストに前記第2の番号に対応する書き込むべきデータを送信するステップとをさらに含む。

40

【0015】

第3の態様では、本発明の実施形態は、ストレージデバイスを提供し、ストレージデバイスは、ホストによって送信される第1のデータ書き込み要求を受信するように構成される受信モジュールであって、前記第1のデータ書き込み要求は、書き込むべきデータ及びアドレス情報を伝送する、受信モジュールと、書き込むべきデータ及びアドレス情報に第1の番号を付与するとともに、第1の番号が付与された前記書き込むべきデータ及び前記アドレス情報をキャッシュに書き込み、第1の番号に対応する前記書き込むべきデータ及び前記アドレス情報を、前記キャッシュから読み出すように構成される読み出し及び書き込みモジ

50

ジュールであって、前記第1の番号は、カレント時間間隔番号である、読出し及び書込みモジュールと、後続のデータ書込み要求において伝送される情報を識別するために、カレント時間間隔番号を変更するように構成されるカレント時間間隔番号マネージャと、前記書き込むべきデータ及び前記アドレス情報を第2のストレージデバイスに送信するように構成される送信モジュールと、を具備する。

【0016】

本発明の実施形態の第3の態様の第1の可能な実装では、前記第1の番号は、現在の複製タスクを識別するために使用され、前記カレント時間間隔番号マネージャは、第2の番号を記録するようにさらに構成され、前記第2の番号は、前記現在の複製タスクの前の最新の完了した複製タスクに対応する番号である。

10

【0017】

本発明の実施形態の第3の態様の第1の可能な実装を参照すると、第3の態様の第2の可能な実装では、前記読出し及び書込みモジュールは、前記第2の番号の後であるとともに、前記第1の番号に先行する番号に対応する書き込むべきデータ及びアドレス情報を、前記キャッシュから読み出すようにさらに構成され、前記送信モジュールは、前記第2の番号の後であるとともに、前記第1の番号に先行する番号に対応する書き込むべきデータ及びアドレス情報を、前記第2のストレージデバイスに送信するようにさらに構成される。

【0018】

本発明の実施形態の第3の態様の第3の実装では、前記カレント時間間隔番号マネージャは、前記カレント時間間隔番号を記録するようにさらに構成され、前記カレント時間間隔番号は、前記第1の番号を生成するために使用される。

20

【0019】

第4の態様では、本発明の実施形態はストレージデバイスを提供し、ストレージデバイスは、第1のストレージデバイスによって送信されたアドレス情報を受信するように構成される受信モジュールと、前記第1のストレージデバイスが故障していることが決定すると、第1の番号に対応する書き込むべきデータを、前記アドレス情報に従って取得するように構成される探索モジュールであって、前記第1の番号に対応するアドレス情報が、前記受信したアドレス情報と同じであるとともに、前記第1の番号がカレント時間間隔番号に先行する番号である、探索モジュールと、前記第1の番号に対応する前記書き込むべきデータ及び前記アドレス情報に第2の番号を付与するとともに、情報をキャッシュに書き込むように構成される書込みモジュールとを具備する。

30

【0020】

本発明の実施形態の第4の態様の第1の可能な実装では、ストレージデバイスは、前記カレント時間間隔番号を記録するように構成されるカレント時間間隔番号マネージャをさらに具備し、前記カレント時間間隔番号は、前記第2の番号を生成するために使用される。

【0021】

本発明の実施形態の第4の態様の第2の可能な実装では、前記受信モジュールは、ホストによって送信されたデータ読出し要求を受信するようにさらに構成され、前記データ読出し要求は、前記受信したアドレス情報を含み、前記探索モジュールは、前記受信したアドレス情報に対応する最新の番号が前記第2の番号であることを決定するようにさらに構成され、前記ストレージデバイスは、送信モジュールをさらに具備し、前記送信モジュールは、前記第2の番号に対応する書き込むべきデータを前記ホストに送信するように構成される。

40

【0022】

第5の態様では、本発明の実施形態は、ストレージデバイスを提供し、ストレージデバイスは、プロセッサと、メモリと、通信バスとを具備し、前記プロセッサと前記メモリは、前記通信バスを使用することによって通信し、前記メモリは、プログラムを格納するように構成され、前記プロセッサは、ホストによって送信される第1のデータ書込み要求を受信するステップであって、前記第1のデータ書込み要求は、書き込むべきデータ及びアドレス情報を伝送する、ステップと、書き込むべきデータ及びアドレス情報に第1の番号

50

を付与するとともに、第1の番号が付与された前記書き込むべきデータ及び前記アドレス情報をキャッシュに書き込むステップであって、前記第1の番号は、カレント時間間隔番号である、ステップと、第1の番号に対応する前記書き込むべきデータ及び前記アドレス情報を、前記キャッシュから読み出すステップと、後続のデータ書き込み要求において伝送される情報を識別するために、カレント時間間隔番号を変更するステップと、前記書き込むべきデータ及び前記アドレス情報を第2のストレージデバイスに送信するステップとを実施する前記プログラムを実行するように構成される。

【0023】

本発明の実施形態の第5の態様の第1の可能な実装では、前記第1の番号は、現在の複製タスクを識別するために使用され、前記プロセッサは、第2の番号を記録するようにさらに構成され、前記第2の番号は、前記現在の複製タスクの前の最新の完了した複製タスクに対応する番号である。

10

【0024】

本発明の実施形態の第5の態様の第1の可能な実装を参照して、第5の態様の第2の可能な実装では、前記プロセッサは、前記第2の番号の後であるとともに、前記第1の番号に先行する番号に対応する書き込むべきデータ及びアドレス情報を、前記キャッシュから読み出し、前記第2の番号の後であるとともに、前記第1の番号に先行する前記番号に対応する前記書き込むべきデータ及び前記アドレス情報を前記第2のストレージデバイスに送信するようにさらに構成される。

【0025】

20

本発明の実施形態の第5の態様の第3の可能な実装では、前記プロセッサは、前記カレント時間間隔番号を記録するようにさらに構成され、前記カレント時間間隔番号は、前記第1の番号を生成するために使用される。

【0026】

第6の様態では、本発明の実施形態はストレージデバイスを提供し、ストレージデバイスは、プロセッサと、メモリと、通信バスとを具備し、前記プロセッサ及び前記メモリは、前記通信バスを使用することによって通信し、前記メモリは、プログラムを格納するように構成され、前記プロセッサは、第1のストレージデバイスによって送信されたアドレス情報を受信するステップと、前記第1のストレージデバイスが故障していることが決定すると、前記アドレス情報に従って前記第2のストレージデバイスによって、第1の番号に対応する書き込むべきデータを取得するステップであって、前記第1の番号に対応するアドレス情報が、前記受信したアドレス情報と同じであるとともに、前記第1の番号がカレント時間間隔番号に先行する番号である、ステップと、前記第1の番号に対応する前記書き込むべきデータ及び前記アドレス情報に第2の番号を付与するとともに、情報をキャッシュに書き込むステップとを実施する前記プログラムを実行するように構成される。

30

【0027】

本発明の実施形態の第6の態様の第1の可能な実装では、前記プロセッサは、前記カレント時間間隔番号を記録するようにさらに構成され、前記カレント時間間隔番号は、前記第2の番号を生成するために使用される。

【0028】

40

本発明の実施形態の第6の態様の第2の可能な実装では、前記プロセッサは、ホストによって送信されたデータ読み出し要求を受信するようにさらに構成され、前記データ読み出し要求は、前記受信したアドレス情報を有し、前記プロセッサは、前記受信したアドレス情報に対応する最新の番号が前記第2の番号であることを決定し、前記ホストに前記第2の番号に対応する書き込むべきデータを送信するようにさらに構成される。

【0029】

本発明の実施形態では、第1のストレージデバイスがホストによって送信されたデータ書き込み要求を受信した後、データ書き込み要求において伝送される情報は、書き込むべきデータ及びアドレス情報を含み、第1のストレージデバイスが、書き込むべきデータ及びアドレス情報に第1の番号を付与するとともに、その情報をキャッシュに書き込み、前記第1

50

の番号はカレント時間間隔番号である。複製タスクがトリガされるとき、第1のストレージデバイスは、キャッシュから第1の番号に対応する書き込むべきデータ及びアドレス情報を読み出し、第2のストレージデバイスにその情報を送信する。加えて、複製タスクがトリガされるとき、第1のストレージデバイスは、カレント時間間隔番号を変更し、その結果、第1のストレージデバイスが、後続のデータ書き込み要求を受信するとき、変更されたカレント時間間隔番号と同じ番号をそのデータ書き込み要求において伝送される情報に付与する。それによって、キャッシュ内では、第2のストレージデバイスに送信されるべきデータ書き込み要求において伝送される情報は、第1のストレージデバイスによって受信されるデータ書き込み要求において伝送される情報と区別される。このことは、キャッシュから第2のストレージデバイスへのデータ書き込み要求において伝送される情報の直接送信を実施する。情報が、データボリュームから読み出す必要なくキャッシュから直接送信されるので、データ複製時間が短く、データ複製の効率が上がる。

10

【0030】

本発明又は従来技術の実施形態の技術的解決策をより明確に説明するために、実施形態又は従来技術を記述することに必要とされる添付の図面を以下に簡潔に紹介する。以下の記述にある添付の図面は、本発明のいくつかの実施形態を示し、当業者は、創造的な努力なしにこれらの図面から別の図面をさらに導くことができる。

【図面の簡単な説明】

【0031】

【図1】本発明の実施形態によるデータ送信方法のアプリケーションネットワークアーキテクチャの概略図である。

20

【図2】本発明の実施形態によるデータ送信方法のフローチャートである。

【図3】本発明の実施形態によるデータ受信方法のフローチャートである。

【図4】本発明の実施形態によるデータ送信方法のシグナリングのフローチャートである。

【図5】本発明の実施形態によるストレージデバイスの概略の構造図である。

【図6】本発明の実施形態による別のストレージデバイスの概略の構造図である。

【図7】本発明の実施形態によるさらに別のストレージデバイスの概略の構造図である。

【図8】本発明の実施形態によるさらに別のストレージデバイスの概略の構造図である。

【図9】本発明の実施形態によるデータ送信方法の別のアプリケーションネットワークアーキテクチャの概略図である。

30

【図10】本発明の実施形態によるデータ複製方法のフローチャートである。

【図11】本発明の実施形態による別のストレージデバイスの概略の構造図である。

【図12】本発明の実施形態によるさらに別のストレージデバイスの概略の構造図である。

【発明を実施するための形態】

【0032】

本発明の実施形態の目的、技術的解決策、及び利点をより分かりやすくする目的で、本発明の実施形態の添付の図面を参照して、本発明の実施形態における技術的解決策を以下に明確に完全に記載する。記載された実施形態は、本発明の実施形態の全てではなく単なる一部である。創造的な努力なしに本発明の実施形態に基づいて当業者によって取得される他の全ての実施形態は、本発明の保護的範囲に含まれるべきである。

40

【0033】

本発明の実施形態中に提供されるデータ送信方法は、ストレージデバイス中に実施され得る。図1は、本発明の実施形態において提供されるデータ送信方法のシステムアーキテクチャの概略図である。図1に示されるように、生産センタは、1つ又は複数の生産ホスト、接続デバイス、及び生産アレイ(以下の実施形態の第1のストレージデバイスに対応する)を含み、障害復旧センタのシステムアーキテクチャは、生産センタのシステムアーキテクチャに類似しており、障害復旧ホスト、接続デバイス、及び障害復旧アレイ(以下の実施形態の第2のストレージデバイスに対応する)を含む。本発明の実施形態では、1つ又

50

は複数の障害復旧センタがあり得る。生産センタ及び障害復旧センタは、IP(Internet Protocol)又はFC(Fiber Channel)を介してデータを転送し得る。生産センタと障害復旧センタの間に制御センタがあり得る。制御センタは、生産センタ側又は障害復旧センタ側に配置されてよく、又は、生産センタと障害復旧センタの間のサードパーティデバイス上に配置されてもよい。生産アレイが故障するとき、制御センタは障害復旧アレイに信号を送信するように構成され、その結果、障害復旧アレイは生産アレイのホストサービスを引き継ぐ。

【0034】

生産ホスト及び障害復旧ホストは、サーバ及びデスクトップコンピュータなど、従来技術で既知である任意のコンピュータデバイスであり得る。オペレーティングシステム、及び他のアプリケーションプログラムが、ホスト内部にインストールされる。

10

【0035】

接続デバイスは、ファイバスイッチ又は他の現在利用可能なスイッチなど、ストレージデバイスとホストの間に、従来技術で既知である任意のインタフェースを含み得る。

【0036】

生産アレイ及び障害復旧アレイは、1つ又は複数の相互接続したRAID(Redundant Array of Inexpensive Disks)、JBOD(Just a Bunch Of Disks)、及びテープライブラリなどの1つ又は複数の相互接続したディスクドライブによって形成されたディスクドライブなどの、従来技術で既知であるストレージデバイスであるか、又は、1つ又は複数の相互接続したディスクドライブがDASD(Direct Access Storage Device)のものである、1つ又は複数の

20

【0037】

生産アレイのストレージ領域は、複数のデータボリュームを含み得る。データボリュームは、物理ストレージ領域をマッピングすることによって形成された論理ストレージ領域である。例えば、データボリュームは、LUN(Logic Unit Number)又はファイルシステムであり得る。本発明のこの実施形態では、障害復旧アレイは生産アレイに類似した構造を有する。

【0038】

図1を参照すると、図1では、本発明の実施形態によるデータ送信方法を説明する。本発明のこの実施形態は、第1のストレージデバイスに適用され、第1のストレージデバイスは、コントローラ、キャッシュメモリ(以降では、キャッシュと称する)、及びストレージ媒体を含む。コントローラは、第1のストレージデバイスのプロセッサであるとともに、IOコマンド及び他のデータサービスを実行するように構成される。キャッシュは、コントローラとハードディスクの間に配置されたメモリであり、少量の容量であるが、ハードディスクより非常に高速である。ストレージ媒体は、第1のストレージデバイスの主メモリであり、一般に磁気ディスクなどの不揮発性ストレージ媒体である。発明の実施形態では、第1のストレージデバイスに含まれる全ての物理ストレージ領域は、ストレージ媒体として参照される。具体的に、以下のステップは、第1のストレージデバイス内のコントローラによって実行され得る。

30

【0039】

ステップS101：第1のストレージデバイスは、ホストによって送信された第1のデータ書き込み要求を受信し、第1のデータ書き込み要求は、書き込むためのデータ及びアドレス情報を伝送する。

40

【0040】

アドレス情報は、LBA(a Logic Block Address)を含み得る。第1のストレージデバイスが複数のデータボリュームを含むとき、アドレス情報は、第1のストレージデバイスのデータボリュームのIDをさらに含み得る。

【0041】

ステップS102：第1の番号を書き込むべきデータ及びアドレス情報に付与し、書き込むべきデータ及びアドレス情報をキャッシュに書き込む。ここで、第1の番号はカレント時

50

間隔番号である。第1のストレージデバイスは、カレント時間間隔番号マネージャを含み得、カレント時間間隔番号マネージャはカレント時間間隔番号を格納する。カレント時間間隔番号は、限定しないが、0、1、2などの数字、又は、a、b、cなどの文字であり得る。

【0042】

第1のデータ書込み要求が受信されると、第1の番号が、第1のデータ書込み要求において伝送される書き込むべきデータとアドレス情報とに付与され、第1の番号は、カレント時間間隔番号によって割り当てられた値である。

【0043】

第1の番号が第1のデータ書込み要求において伝送される情報に付与された後、第1のデータ書込み要求において伝送される変化した情報は、キャッシュに書き込まれ、その結果、第1のデータ書込み要求において伝送される書き込むべきデータ、アドレス情報、及び第1の番号は、全てキャッシュに格納される。

10

【0044】

加えて、時間間隔内で、別のデータ書込み要求も受信され得、データ書込み要求において伝送される情報に第1の番号を加え、キャッシュにその情報を書き込むことも必要になる。カレント時間間隔番号が変わる前に、第1の番号が全てのデータ書込み要求において伝送される情報に付与されることに留意すべきである。

【0045】

ステップS103：第1の番号に対応する書き込むべきデータ及びアドレス情報をキャッシュから読み出す。

20

【0046】

複製タスクがトリガされると、第1の番号に対応する書き込むべきデータ及びアドレス情報をキャッシュから読み出し得る。第1の番号に対応する書き込むべきデータ及びアドレス情報の複数の部分があり得ることがわかる。

【0047】

複製タスクは、第1のストレージデバイスが第2のストレージデバイスに、時間期間内に1つのデータボリュームによって受信されたデータ書込み要求において伝送される情報を送信することを意味し、カレント時間間隔番号と同じ番号が、全てのデータ書込み要求において伝送される情報に付与される。複製タスクは、限定ではないが、タイマによってトリガされるか、又は手動でトリガされ得る。複製の目的は、第1のストレージデバイスによって受信されたデータ書込み要求において伝送される書き込むべきデータを第2のストレージデバイスに送信することであり、その結果、第2のストレージデバイスは、第1のストレージデバイスが故障したときに、第1のストレージデバイスの働きを引き継ぐことができる。データ書込み要求において伝送されるアドレス情報(LBAなど)は、第2のストレージデバイスに送信される必要があり、LBAは、第2のストレージデバイスが書き込むべきデータを格納するアドレスを示すために使用されることがわかる。第2のストレージデバイスが、第1のストレージデバイスとして同様の物理構造を有するので、第1のストレージデバイスに適用可能なLBAは、第2のストレージデバイスにも適用可能である。

30

【0048】

本発明の実施形態では、複製タスクは、第1のストレージデバイスのデータボリュームに特有である。第1のストレージデバイスが複数のデータボリュームを含むとき、各データボリュームは1つの複製タスクに対応する。

40

【0049】

ステップS104：後続のデータ書込み要求において伝送される情報を識別するために、カレント時間間隔番号を変更する。

【0050】

複製タスクがトリガされるとき、カレント時間間隔番号マネージャが、カレント時間間隔番号を変更する必要がある。後続のデータ書込み要求が受信されると、別の番号が、後続のデータ書込み要求において伝送される情報に付与される必要があり、別の番号は、変

50

更されたカレント時間間隔番号によって割り当てられた値である。それによって、キャッシュ内では、第2のストレージデバイスに送信されるデータ書込み要求において伝送された情報が、第1のストレージデバイスによって受信されるデータ書込み要求において伝送される情報と区別され得る。

【0051】

ステップS103及びステップS104が時系列に依存しないことに留意すべきである。

ステップS105：書き込むべきデータ及びアドレス情報を第2のストレージデバイスに送信する。

【0052】

第1のストレージデバイスは、第1の番号に対応するとともにキャッシュから読み出される書き込むべきデータ及びアドレス情報を第2のストレージデバイスに送信する。

10

【0053】

具体的には、第1のストレージデバイスは、読み出された全ての書き込むべきデータ及びアドレス情報を第2のストレージデバイスに直接送信し得る。すなわち、第1のストレージデバイスは、第2のストレージデバイスのデータボリュームのIDを取得した後、各データ書込み要求において伝送される書き込むべきデータ及びアドレス情報に従って新しいデータ書込み要求と、第2のストレージデバイスのデータボリュームのIDとを生成し、第2のストレージデバイスに新しいデータ書込み要求を送信する。

【0054】

本発明の実施形態では、第1のストレージデバイスが、ホストによって送信されたデータ書込み要求を受信した後、データ書込み要求において伝送される情報は、書き込むべきデータ及びアドレス情報を含み、第1のストレージデバイスは、書き込むべきデータ及びアドレス情報に第1の番号を付与し、キャッシュにその情報を書き込む。第1の番号は、カレント時間間隔番号である。複製タスクがトリガされると、第1のストレージデバイスは、第1の番号に対応する書き込むべきデータ及びアドレス情報をキャッシュから読み出し、第2のストレージデバイスに情報を送信する。加えて、複製タスクがトリガされると、第1のストレージデバイスはカレント時間間隔番号を変更し、その結果、第1のストレージデバイスは、後続のデータ書込み要求を受信するとき、データ書込み要求において伝送される情報の変更されたカレント時間間隔番号と同じ番号を付与する。それによって、キャッシュ内では、第2のストレージデバイスに送信されるデータ書込み要求において伝送される情報は、第1のストレージデバイスによって受信されるデータ書込み要求において伝送される情報と区別される。このことは、キャッシュから第2のストレージデバイスへのデータ書込み要求において伝送される情報の直接送信を実施する。情報は、データボリュームから読み出す必要なくキャッシュから直接送信されるので、データ複製時間が短く、それよりデータ複製の効率が上がる。上述の実施形態では、複製タスクがトリガされるとき、第1のストレージデバイスは、カレント時間間隔番号に対応する書き込むべきデータ及びアドレス情報を第2のストレージデバイスに送信するとともに、後続のデータ書込み要求において伝送される情報を識別するために、カレント時間間隔番号を変更することがわかる。

20

30

【0055】

後続の複製タスクがトリガされるとき、第1のストレージデバイスは、第2のストレージデバイスに変更したカレント時間間隔番号に対応する書き込むべきデータ及びアドレス情報を送信するとともに、カレント時間間隔番号を再度変更する。このことによって、第1のストレージデバイスが、完全に受信したデータ書込み要求において伝送される情報をまとめて第2のストレージデバイスに送信することを保証する。

40

【0056】

しかしながら、複数の障害復旧センタがあるとき、第2の障害復旧センタに対応するストレージデバイスが第3のストレージデバイスであると仮定すると、第1のストレージデバイスはまた、第1のストレージデバイスによって受信したデータ書込み要求において伝送される情報を第3のストレージデバイスに送信する必要がある。第2のストレージデバイス

50

のために、複製タスクがトリガされると、カレント時間間隔番号マネージャは、カレント時間間隔番号を変更する。この時点で、第2のストレージデバイスにカレント時間間隔番号によって割り当てられるとともに、第3のストレージデバイスに割り当てられる番号は、両方とも変更された番号である。しかしながら、カレント時間間隔番号が変更される前にその番号に対応するデータ書込み要求において伝送される情報は、まだ第3のストレージデバイスに送信されていない。

【0057】

従って、複数の障害復旧センタのシナリオでは、上述の実施形態はさらに以下のステップを含み得る。

【0058】

ステップS106：第2の番号を記録する。第2の番号は、現在の複製タスクの前の、最新の完了した複製タスクに対応する番号である。上述の実施形態では、第1の番号は、カレント時間間隔番号と同じであり、現在の複製タスクを識別するために使用され得る。現在の複製タスクは、第1のストレージデバイスが、時間の現在の期間内に、1つのデータボリュームによって受信されたデータ書込み要求において伝送される情報を第2のストレージデバイスに送信することを意味し、カレント時間間隔番号と同じ番号が、全てのデータ書込み要求において伝送される情報に付与される。

【0059】

第2の番号は、現在の複製タスクの前の最新の完了した複製タスクに対応する番号である。

【0060】

複数の障害復旧センタがあるとき、カレント時間間隔番号は、複製タスクが別の障害復旧センタ内のストレージデバイスに開始されるときに変更され得る。従って、以前に完了した複製タスクに対応する番号を記録する必要がある。

【0061】

別の番号が第2の番号と第1の番号の間に存在する場合、この番号に対応するデータ書込み要求において伝送される情報は、第2のストレージデバイスに送信されず、ステップS107が実行される必要がある。

【0062】

ステップS107：第2の番号に後続するとともに第1の番号に先行する番号に対応する書き込むべきデータ及びアドレス情報をキャッシュから読み出す。

【0063】

特定の読み出しプロセスはステップS103と同様であり、本明細書にはさらに記載しない。

ステップS107及びステップS103は、時系列に依存せず、同時に実行されてよいことに留意すべきである。

【0064】

ステップS108：第2の番号に後続するとともに第1の番号に先行する番号に対応する書き込むべきデータ及びアドレス情報を送信する。

【0065】

特定の送信プロセスは、ステップS105と同様であり、本明細書にはさらに記載しない。本発明の実施形態では、カレント時間間隔番号に対応するデータ書込み要求において伝送される情報が、第2のストレージデバイスに送信されるだけでなく、以前に完了した複製タスクに対応する番号と現在の時間期間番号の間の番号に対応するデータ書込み要求において伝送される情報も、第2のストレージデバイスに送信される。このことは、複数の障害復旧センタのシナリオに適用可能であり、データ複製の完全性を保証する。

【0066】

図2を参照すると、図2では、本発明によるデータ受信方法の実施形態を説明する。本発明の実施形態は、障害復旧センタが生産センタによって送信されたデータ書込み要求において伝送される情報を受信するというシナリオに適用可能である。方法は、以下のステッ

10

20

30

40

50

プを含み得る。

【0067】

ステップS201：第2のストレージデバイスが、第1のストレージデバイスによって送信されたアドレス情報を受信する。

【0068】

具体的には、第2のストレージデバイスは、第1のストレージデバイスによって送信された書き込むべきデータ及びアドレス情報を受信し得るか、又は第2のストレージデバイスは、第1のストレージデバイスによって送信されたデータ書き込み要求を受信し得、データ書き込み要求は、書き込むべきデータ及びアドレス情報を含み、アドレス情報は、LBAであってよい。第2のストレージデバイスが複数のデータボリュームを含むとき、アドレス情報は、第2のストレージデバイスのデータボリュームのIDをさらに含み得る。アドレス情報が1つ以上あり得ることがわかる。

10

【0069】

書き込むべきデータ及びアドレス情報を受信した後、第2のストレージデバイスは、書き込むべきデータ及びアドレス情報にカレント時間間隔番号と同じ番号を付与し、その情報をキャッシュに書き込む。その結果、カレント時間間隔番号と同じ番号、書き込むべきデータ、及びアドレス情報がキャッシュに格納される。

【0070】

第2のストレージデバイスはまた、カレント時間間隔番号マネージャを含むこと、及びカレント時間間隔番号マネージャがカレント時間間隔番号を格納することに留意すべきである。限定しないが、カレント時間間隔番号は、0、1、2などの数字、又は、a、b、cなどの文字であり得る。ここでのカレント時間間隔番号は、第1のストレージデバイスのカレント時間間隔番号と関係なくてよい。

20

【0071】

ステップS202：第1のストレージデバイスが故障していることが決定すると、第2のストレージデバイスは、アドレス情報に従って第1の番号に対応する書き込むべきデータを取得する。第1の番号に対応するアドレス情報は、受信したアドレス情報と同じであり、第1の番号は、カレント時間間隔番号に先行する番号である。

【0072】

一般に、第1のストレージデバイス及び第2のストレージデバイスの両方が正常に動作する場合、第2のストレージデバイスは、第1のストレージデバイスによって送信された全てのデータ書き込み要求において伝送される情報を受信し、各データ書き込み要求において伝送される情報に、カレント時間間隔番号と同じ番号を付与し、キャッシュにその情報を格納し得る。しかしながら、第1のストレージデバイスが故障した場合、第2のストレージデバイスは、第1のストレージデバイスのカレント時間間隔番号に対応する書き込むべきデータの一部のみを受信することがある。この場合、第2のストレージデバイスによって格納されたデータは、不正確である場合があり、かつ第2のストレージデバイスが直接第1のストレージデバイスの働きを引き継ぐ場合、データの一貫性が保証され得ない。例えば、ホストが、この時点で第2のストレージデバイスに読み出しデータ要求を送信する場合、アドレス情報に格納されたデータを読み出すことを要求するとき、第2のストレージデバイスがアドレス情報に対応する最新の番号を探索し、その後カレント時間間隔番号に対応する書き込むべきデータ（しかしながら、このデータは不正確である）をホストに送信する。従って、この場合、第2のストレージデバイスのキャッシュ内のデータから第2のストレージデバイスのカレント時間間隔番号に先行する番号に対応するデータをリストアする必要がある。

30

40

【0073】

具体的には、第1のストレージデバイスが故障していることは、制御センタが第2のストレージデバイスに信号を送信する方法によって判断され得、信号は、第1のストレージデバイスが故障していること、及び第2のストレージデバイスが第1のストレージデバイスのホストサービスを引き継ぐ必要があることを示すために使用される。

50

【 0 0 7 4 】

一般に、複製タスクが完了するとき、制御センタが第1のストレージデバイス及び第2のストレージデバイスの両方に複製成功指示を送信し得る。第2のストレージデバイスが指示を受信しない場合、現在の複製タスクが完了していないことを示す。複製タスクの完了は、第1のストレージデバイスがカレント時間間隔番号に対応する全てのデータ書込み要求において伝送される情報を第2のストレージデバイスに送信したこと、及び第2のストレージデバイスがその情報を受信し終わったことを意味する。

【 0 0 7 5 】

第2のストレージデバイスが、第1のストレージデバイスが故障していることが決定すると、現在の複製タスクが完了する場合、第2のストレージデバイスは、直接第1のストレージデバイスの働きを引き継ぐことができ、データの一貫性が保証され得る。この状況は、本発明の実施形態の議論の範囲外である。

10

【 0 0 7 6 】

しかしながら、現在の複製タスクが完了しない場合、第2のストレージデバイスのキャッシュ内のデータから第2のストレージデバイスのカレント時間間隔番号に先行する番号に対応するデータをリストアする必要がある。

【 0 0 7 7 】

特定のリストア方法は、受信したアドレス情報に従って、カレント時間断片番号に先行する番号に対応するアドレス情報が受信したアドレス情報と同一であるか否かを検査し、同一でない場合、受信したアドレス情報と同一のアドレス情報が見つかるまでさらに先行する番号に対応するアドレス情報を検査することを継続し、見つかったアドレス情報の番号に対応する書き込むべきデータを取得し得る。

20

【 0 0 7 8 】

ステップS203：第1の番号に対応する書き込むべきデータ及びアドレス情報に第2の番号を付与し、その情報をキャッシュに書き込む。

【 0 0 7 9 】

第2の番号は、カレント時間間隔番号を変更することによって取得された番号であるとともに、本発明の実施形態ではキャッシュに格納された最新の番号でもある。アドレス情報内に格納されたデータを読み出すことを要求するために、ホストが第2のストレージデバイスにデータ読出し要求を送信するとき、第2のストレージデバイスは、アドレス情報に対応する最新の番号が第2の番号であることを探索することによって学習し、第2の番号に対応する書き込むべきデータをホストに送信する。それによって、データの一貫性が保証される。

30

【 0 0 8 0 】

本発明の実施形態では、第2のストレージデバイスは、第1のストレージデバイスによって送信されたアドレス情報を受信し、第1のストレージデバイスが故障するとき、カレント時間間隔番号に先行する番号に対応する書き込むべきデータを取得し、カレント時間間隔番号に先行する番号に対応する書き込むべきデータ及びアドレス情報に第2の番号を付与し、キャッシュにその情報を格納する。それによって、データの一貫性が保証される。

【 0 0 8 1 】

図3を参照すると、図3では、本発明によるデータ送信方法の実施形態を説明する。本発明の実施形態では、生産アレイ内のキャッシュと障害復旧アレイ内のキャッシュとを区別するために、生産アレイ内のキャッシュが第1のキャッシュとして参照され、障害復旧アレイ内のキャッシュが第2のキャッシュとして参照される。

40

【 0 0 8 2 】

図3に示されるように、方法は、以下のステップを含む。

ステップS301：生産アレイは、生産ホストからデータ書込み要求Aを受信する。データ書込み要求Aは、ボリュームID、書込みアドレスA、及び書き込むべきデータAを含む。書込みアドレスAは、LBAなど、書き込むべきデータAが書き込まれる生産アレイの論理アドレスである。一般に、生産アレイが、データ書込み要求Aを実行するとき、生産アレイは

50

、LBAをPBA(Physics Block Address)に変換する必要があり、その後PBAに従ってストレージ媒体に書き込むべきデータAを書き込む。ボリュームIDは、データ書き込み要求Aに対応するデータボリュームのIDである。実施形態では、生産アレイは、1つのボリューム(以下、プライマリボリュームと称する)を含むと仮定する。データ書き込み要求Aにおいて伝送される情報は、プライマリボリュームID、書き込みアドレスA、及び書き込みアドレスAを含む。

【0083】

ステップS302：生産アレイは、データ書き込み要求Aをデータ書き込み要求A'に変更し、データ書き込み要求A'は、データ書き込み要求Aにおいて伝送される情報と第1の番号を含む。

【0084】

本発明の実施形態では、生産アレイのコントローラは、カレント時間間隔番号(CTPN)マネージャを含み得る。CTPNマネージャは、第1の番号を生成するために使用されるカレント時間間隔番号を記録する。具体的には、第1の番号はカレント時間間隔番号に等しい。

【0085】

データ書き込み要求Aを受信した後、生産アレイは、データ書き込み要求Aをデータ書き込み要求A'に変更する。具体的には、変更方法は、データ書き込み要求Aにおいて伝送される情報に第1の番号を付与することであり得る。例えば、カレント時間間隔番号は1であり、第1の番号も1である。

【0086】

選択的に、書き込むべきデータAが受信され、タイムスタンプに対応する番号を判断するために、事前に格納された番号列内でタイムスタンプが合致するとき、タイムスタンプが記録され得る。具体的には、限定ではないが、番号列は、マッピングテーブル又は他の形式であり得る。表1に示されるように、番号列は複数の番号を含み、各番号はタイムスタンプの区間に対応する。

【0087】

【表1】

表 1

番号	タイムスタンプの セクション
1	9:01-10:00
2	10:01-11:00
3	11:01-12:00

【0088】

データ書き込み要求Aが受信されたときのタイムスタンプが9時30分であると仮定すると、対応する番号は1であり、データ書き込み要求Aは、番号に従ってデータ書き込み要求A'に変更され得る。

【0089】

ステップS303：生産アレイは、データ書き込み要求A'を第1のキャッシュに書き込み、その結果、データ書き込み要求A'において伝送される情報は、第1のキャッシュに格納される。データ書き込み要求A'において伝送される情報は、第1の番号、プライマリボリュームID、書き込みアドレスA、及び書き込むべきデータAを含む。

【0090】

本発明の実施形態では、第1の番号は、複数のデータ書込み要求に対応し得る。CTPN内に記録されたカレント時間間隔番号が変更される前に、第1の番号が全ての受信されたデータ書込み要求において伝送される情報に付与される。

【0091】

データ書込み要求Aが受信された後、データ書込み要求Bが受信されるとともにデータ書込み要求B'に変更され得、その結果、データ書込み要求B'は第1の番号をさらに含み、データ書込み要求Cも受信されるとともにデータ書込み要求C'に変更され得、その結果、データ書込み要求C'は、第1の番号をさらに含むことがわかる。

【0092】

例えば、データ書込み要求A'、データ書込み要求B'、及びデータ書込み要求C'が第1のキャッシュに書き込まれた後、第1のキャッシュに書き込まれた情報は、表2に示さ

10

【0093】

【表2】

表 2

番号	ボリューム ID	書込みアドレス	書き込むべきデータ
1	プライマリボリューム ID	書込みアドレス A	書き込むべきデータ A
1	プライマリボリューム ID	書込みアドレス B	書き込むべきデータ B
1	プライマリボリューム ID	書込みアドレス C	書き込むべきデータ C

20

30

【0094】

本発明の実施形態では、生産アレイが1つのデータボリューム(プライマリボリュームとして参照され得る)を含むと仮定すると、データ書込み要求A'、データ書込み要求B'、及びデータ書込み要求C'において伝送されるデータボリュームのIDは、全てプライマリボリュームIDであることに留意すべきである。本発明の別の実施形態では、生産アレイは、複数のデータボリュームを含み得、従って、データ書込み要求A'、データ書込み要求B'、及びデータ書込み要求C'において伝送されるデータボリュームのIDは、異なり得る。加えて、表2は、第1のキャッシュのデータ書込み要求において伝送される情報のストレージの例にすぎない。代替として、情報がツリー形式で格納されてよい。ストレージの形態は、本明細書では限定されない。

40

【0095】

例として表2を取り上げると、番号、ボリュームID、及び書込みアドレスは、表2のインデックスとみなされ得る。インデックスに従って、対応する書き込むべきデータを見つけることができる。インデックスが同じとき、インデックスに対応する書き込むべきデータも、同じあるべきである。従って、新しいデータ書込み要求が書き込まれると、番号、ボ

50

リユームID、及び新しいデータ書込み要求の書込みアドレスと同じ情報が第1のキャッシュに格納されるか否かを判断する必要がある、格納される場合、新しいデータ書込み要求において伝送される情報が、古い情報を上書きするために使用される。データ書込み要求A'、データ書込み要求B'、及びデータ書込み要求C'が第1のキャッシュに書き込まれるとき、それらの番号、ボリュームID、及び書込みアドレスが第1のキャッシュにすでに格納された情報と同じであるか否かを判断する必要もあり、それらの番号、ボリュームID、及び書込みアドレスが、第1のキャッシュ内に既に格納された情報と異なるので、データ書込み要求A'、データ書込み要求B'、及びデータ書込み要求C'は、第1のキャッシュに全て書き込まれ得ることがわかる。

【0096】

例えば、データ書込み要求Dが受信され、データ書込み要求DがプライマリボリュームID、書込みアドレスB、書き込むべきデータDを含む場合、データ書込み要求Dは、データ書込み要求D'に変更され、その結果、データ書込み要求D'は、第1の番号をさらに含む。従って、データ書込み要求D'が第1のキャッシュに書き込まれるとき、番号、ボリュームID、及びデータ書込み要求D'の書込みアドレスと同じ情報が第1のキャッシュに格納されるか否かを判断する必要がある、格納される場合、データ書込み要求D'において伝送される情報は、古い情報を上書きするために使用される。データ書込み要求D'において伝送される番号、ボリュームID、及び書込みアドレスは、第1のキャッシュ内の、データ書込み要求B'内に含まれる番号、ボリュームID、及び書込みアドレスと全て同じであるので、データ書込み要求D'の情報は、データ書込み要求B'の情報を上書きする。

【0097】

具体的には、データ書込み要求D'が第1のキャッシュに書き込まれた後、第1のキャッシュに格納された情報は表3に示され得る。

【0098】

【表3】

表3

番号	ボリューム ID	書込みアドレス	書き込むべきデータ
1	プライマリボリューム ID	書込みアドレス A	書き込むべきデータ A
1	プライマリボリューム ID	書込みアドレス B	書き込むべきデータ D
1	プライマリボリューム ID	書込みアドレス C	書き込むべきデータ C

【0099】

ステップS304：複製タスクがトリガされるとき、生産アレイは、CTPNマネージャ内に含まれるカレント時間間隔番号を変更し、例えば、1から2にカレント時間間隔番号を変更し得る。

【0100】

生産アレイのカレント時間間隔番号と障害復旧アレイのカレント時間間隔番号とを区別するために、本発明の実施形態では、生産アレイのカレント時間間隔番号は第1のカレント時間間隔番号として参照され、障害復旧アレイのカレント時間間隔番号は、第2のカレント時間間隔番号として参照される。

【0101】

第1のカレント時間間隔番号が1から2に変更された後、それに応じて、番号2は、後に受信される全てのデータ書込み要求において伝送される情報に付与されることがわかる。例えば、データ書込み要求Eが受信され、データ書込み要求EがプライマリボリュームID、書き込むべきデータアドレスA、及び書き込むべきデータEを含み、データ書込み要求Eがデータ書込み要求E'に変更され、その結果、データ書込み要求E'がさらに番号2を含む。デ

ータ書込み要求Fが受信され、データ書込み要求Fが、プライマリボリュームID、書込みアドレスF、及び書き込むべきデータFを含み、書き込むべきデータFが書き込むべきデータF'に変更され、その結果、データ書込み要求F'はさらに番号2を含む。データ書込み要求E'及びデータ書込み要求F'は、第1のキャッシュに書き込まれ、第1のキャッシュに格納された情報は、表4に示され得る。

【0102】

【表4】

表4

番号	ボリュームID	書込みアドレス	書き込むべきデータ
1	プライマリボリュームID	書込みアドレスA	書き込むべきデータA
1	プライマリボリュームID	書込みアドレスB	書き込むべきデータD
1	プライマリボリュームID	書込みアドレスC	書き込むべきデータC
2	プライマリボリュームID	書込みアドレスA	書き込むべきデータE
2	プライマリボリュームID	書込みアドレスF	書き込むべきデータF

【0103】

ステップS305：障害復旧アレイは、障害復旧アレイのCTPNマネージャ内に含まれる第2のカレント時間間隔番号を変更し得、例えば11から12に変更し得る。

【0104】

本発明の実施形態では、障害復旧アレイはまた、障害復旧アレイ自身のCPTNマネージャを含んでもよい。生産アレイの複製タスクがトリガされると、生産アレイのCTPNマネージャは第1のカレント時間間隔番号を変更し、制御センタはまた、障害復旧アレイに制御信号を送信し、その結果、障害復旧アレイはまた、そのCPTNマネージャに含まれる第2のカレント時間間隔番号を変更する。従って、ステップS305及びステップS304は、時系列に依存しない。

【0105】

ステップS306A：生産アレイは、第1の番号に対応するデータ書込み要求において伝送される情報を第1のキャッシュから読み出す。

【0106】

具体的には、上記記載に従って、第1の番号に対応するデータ書込み要求において伝送される情報が表3に示される。

【0107】

ステップS306B：生産アレイは、障害復旧アレイに書き込まれるデータボリュームのIDを取得する。

【0108】

ステップS306C：生産アレイは、データボリュームのID、及び第1の番号に対応するデータ書込み要求において伝送される情報に従って、新しいデータ書込み要求を生成する。

【0109】

具体的には、データ書込み要求A'は、データボリュームのID、書込みアドレスA、及び書き込むべきデータAに従って生成され得、データ書込み要求D'は、データボリュームのID、書込みアドレスB、及び書き込むべきデータDに従って生成され得、データ書込み要求C'は、データボリュームのID、書込みアドレスC、及び書き込むべきデータCに従って生成され得る。

【0110】

本発明の別の実施形態では、生産アレイ及び障害復旧アレイの両方は、複数のデータボ

10

20

30

40

50

リユームを含み得、従って、データ書込み要求A'、データ書込み要求D'、及びデータ書込み要求C'に含まれるデータボリュームのIDは異なり得る。しかしながら、障害復旧アレイ内のデータボリュームのIDが、生産アレイ内のデータボリュームのIDに1対1でマッピングされる。

【0111】

ステップS307：生産アレイは、障害復旧アレイに新しいデータ書込み要求を送信する。具体的には、生産アレイは、データ書込み要求A'、データ書込み要求D'、及びデータ書込み要求C'を、障害復旧アレイに送信する。

【0112】

ステップS308：障害復旧アレイは、受信したデータ書込み要求を変更する。
例えば、障害復旧アレイは、CTPNマネージャ内に記録された第2のカレント時間間隔番号に従って、データ書込み要求A'をデータ書込み要求A''に変更し得る。具体的には、変更方法は、データ書込み要求A'において伝送される情報に、番号12を付与することであり得る。

10

【0113】

同様に、番号12は、データ書込み要求B'において伝送される情報に付与され得、データ書込み要求B'がデータ書込み要求B''に変更され、番号12は、データ書込み要求C'において伝送される情報に付与され得、データ書込み要求C'がデータ書込み要求C''に変更される。

【0114】

ステップS309：障害復旧アレイは、変更したデータ書込み要求を第2のキャッシュに書き込む。

20

【0115】

具体的には、第2のキャッシュに格納された情報は、表5に示され得る。

【0116】

【表5】

表 5

番号	ボリューム ID	書込みアドレス	書き込むべきデータ
12	セカンダリボリューム ID	書込みアドレス A	書き込むべきデータ A
12	セカンダリボリューム ID	書込みアドレス B	書き込むべきデータ D
12	セカンダリボリューム ID	書込みアドレス C	書き込むべきデータ C

30

【0117】

ステップS310：データ書込み要求内の書込みアドレスに従って、障害復旧アレイは、書込みアドレスに対応するストレージ媒体に書き込むべきデータを書き込む。

【0118】

一般に、キャッシュ領域は制限される。
従って、キャッシュ領域の利用率が特定のしきい値に達するとき、ハードディスクにキャッシュ内のデータを書き込む必要がある。具体的には、書き込むべきデータAは、書込みアドレスAに対応するストレージ媒体に書き込まれ、書き込むべきデータDは書込みアドレスBに対応するストレージ媒体に書き込まれ、書き込むべきデータCは書込みアドレスCに対応するストレージ媒体に書き込まれる。

40

【0119】

ステップS311：データ書込み要求内の書込みアドレスに従って、生産アレイは、書込みアドレスに対応するストレージ媒体に書き込むべきデータを書き込む。

【0120】

50

同様に、生産アレイのキャッシュ領域の利用率が特定のしきい値に達するとき、ハードディスクにキャッシュ内のデータを書き込む必要がある。上述からわかるように、以下の情報が第1のキャッシュに格納される。

【 0 1 2 1 】

【 数 1 】

番号	ボリューム ID	書込みアドレス	書き込むべきデータ
1	プライマリボリューム ID	書込みアドレス A	書き込むべきデータ A
1	プライマリボリューム ID	書込みアドレス B	書き込むべきデータ D
1	プライマリボリューム ID	書込みアドレス C	書き込むべきデータ C
2	プライマリボリューム ID	書込みアドレス A	書き込むべきデータ E
2	プライマリボリューム ID	書込みアドレス F	書き込むべきデータ F

10

【 0 1 2 2 】

具体的には、同じボリュームID、同じ書込みアドレス、及び異なる番号をもつデータ書込み要求に対して、より小さい番号のデータ書込み要求において伝送される書き込むべきデータが最初に書き込まれ、その後、大きな番号のデータ書込み要求において伝送される書き込むべきデータが書き込まれる。例えば、書き込むべきデータDが最初に書き込まれ、続いて書き込むべきデータEが書き込まれる。又は、より大きな番号をもつデータ書込み要求において伝送される書き込むべきデータが、より小さな番号をもつデータ書込み要求において伝送される書き込むべきデータを書き込むことなく、直接書き込まれる。例えば、書き込むべきデータEが直接書き込まれる。

20

【 0 1 2 3 】

ステップS310及びステップS311は、時系列に依存しない。

ステップS312：複製タスクがトリガされると、生産アレイは、生産アレイのCTPNマネージャに含まれる第1のカレント時間間隔番号を変更し得、例えば、カレント時間間隔番号を2から3に変更し得る。

30

【 0 1 2 4 】

生産アレイのCTPN内の第1のカレント時間間隔番号が2から3に変更された後、それに応じて、後に生産アレイによって受信される全てのデータ書込み要求において伝送される情報に番号3が付与されることがわかる。

【 0 1 2 5 】

ステップS313：障害復旧アレイが障害復旧アレイのCTPNマネージャに含まれる第2のカレント時間間隔番号を変更し得、例えば、第2のカレント時間間隔番号を12から13に変更し得る。

【 0 1 2 6 】

障害復旧アレイのCTPNマネージャ内の第2のカレント時間間隔番号が、12から13に変更された後、それに応じて、後に障害復旧アレイによって受信される全てのデータ書込み要求において伝送される情報に番号13が付与されることがわかる。

40

【 0 1 2 7 】

ステップS314：生産アレイは、番号2に対応するデータ書込み要求において伝送される情報を読み出し、対応するデータ書込み要求を生成し、障害復旧アレイに生成したデータ書込み要求を送信する。

【 0 1 2 8 】

具体的に、上述の記載からわかるように、番号2に対応するデータ書込み要求において伝送される情報は、データ書込み要求Eにおいて伝送される情報とデータ書込み要求Fにおいて伝送される情報とを含む。同様に、生産アレイが障害復旧アレイのデータボリューム

50

のIDを取得した後、生産アレイは、データボリュームのID、書込みアドレスA、及び書き込むべきデータEに従って、データ書込み要求E' 'を生成し得、データボリュームのID、書込みアドレスF、及び書き込むべきデータFに従って、データ書込み要求F' 'を生成し得る。従って、生産アレイによって障害復旧アレイに送信されたデータ書込み要求は、データ書込み要求E' 'とデータ書込み要求F' 'である。

【0129】

本発明の実施形態では、生産アレイは、時系列に障害復旧アレイにデータ書込み要求を送信せず、データ書込み要求をランダムに送信し得ることに留意すべきである。具体的には、生産アレイは、最初にデータ書込み要求E' 'を、その後にデータ書込み要求F' 'を送信してよく、又は、最初にデータ書込み要求F' 'を、その後にデータ書込み要求E' 'を送信してもよい。

10

【0130】

上述からわかるように、この時点で、障害復旧アレイのCTPNマネージャ内の第2のカレント時間間隔番号は13である。従って、データ書込み要求E' 'を受信した後、障害復旧アレイは、番号13を含むデータ書込み要求E' ' 'にデータ書込み要求E' 'を変更する必要がある。同様に、データ書込み要求F' 'を受信した後、障害復旧アレイが番号13を含むデータ書込み要求F' ' 'にデータ書込み要求F' 'を変更する必要がある。

【0131】

ステップS315：障害復旧アレイが生産アレイのホストサービスを引き継ぐための指示を受信する。

20

【0132】

本発明の実施形態では、生産アレイが故障する場合、障害復旧アレイが生産アレイのホストサービスを引き継ぐ必要がある。従って、障害復旧アレイがデータの一貫性の要求を満足する必要がある。

【0133】

ステップS314からわかるように、現在の複製タスクにおいて、障害復旧アレイによって受信されるデータ書込み要求が、データ書込み要求E' '及びデータ書込み要求F' 'を含む。変更されたデータ書込み要求E' '及びデータ書込み要求F' 'が、両方とも第2のキャッシュへの書き込みが成功するまで、障害復旧アレイは生産アレイのホストサービスを引き継がない。障害復旧アレイがホストサービスを引き継ぐことを開始するとき、現在の複製サイクルが完了し、データの一貫性の要求が満たされたことを示す。

30

【0134】

障害復旧アレイがデータ書込み要求E' 'をデータ書込み要求E' ' 'に変更し、データ書込み要求F' ' 'の第2のキャッシュへの書き込みに成功する前に、データ書込み要求E' ' 'の第2のキャッシュへの書き込みが成功した後で、生産アレイが故障し、障害復旧アレイが生産アレイのホストサービスを引き継ぐことを開始する場合、現在の複製タスクが未完了であり、かつデータの一貫性の要求が満足されない。同様に、障害復旧アレイがデータ書込み要求F' 'をデータ書込み要求F' ' 'に変更し、データ書込み要求E' ' 'の第2のキャッシュへの書き込みに成功する前に、データ書込み要求F' ' 'の第2のキャッシュへの書き込みに成功した後で、生産アレイが故障し、障害復旧アレイが生産アレイのホストサービスを引き継ぐことを開始する場合、現在の複製タスクが未完了であり、かつデータの一貫性の要求が満足されない。

40

【0135】

この場合、番号12に対応する複製タスクが完了したときの状態に障害復旧アレイのキャッシュ内のデータをリストアする必要がある。以下では、障害復旧アレイがデータ書込み要求E' 'をデータ書込み要求E' ' 'に変更するとともに、データ書込み要求E' ' 'を第2のキャッシュに成功裏に書き込むこと、かつデータ書込み要求F' ' 'を第2のキャッシュに成功裏に書き込みしないと仮定する。

【0136】

ステップS316：障害復旧アレイが、現在の複製サイクルで第2のキャッシュに書き込み

50

が成功したデータ書込み要求において伝送される書込みアドレスを取得する。

【0137】

上述からわかるように、番号13に対応する複製タスクでは、データ書込み要求E' ' 'の第2のキャッシュへの書き込みが成功し、データ書込み要求E' ' 'において伝送される書込みアドレスは、書込みアドレスAである。

【0138】

ステップS317:書込みアドレスに従って、障害復旧アレイは、書込みアドレスAと同じ書込みアドレスを探索するために、先行する番号に対応するデータ書込み要求において伝送される情報内でマッチングを実行する。

【0139】

その書込みアドレスと同じ書込みアドレスが見つかり、S318が実行される。そうでない場合、マッチングは、データ書込み要求E' ' 'において伝送される書込みアドレスと同じアドレスが見つかるまで、さらに先行する番号(例えば、番号11)に対応するデータ書込み要求において伝送される情報内で実行されることが継続される。

【0140】

上述からわかるように、番号12に対応するデータ書込み要求において伝送される情報は、表5に示される。データ書込み要求A' ' 'において伝送される書込みアドレスは、データ書込み要求E' ' 'において伝送される書込みアドレスと同じである。

【0141】

障害復旧アレイが、データボリュームのIDを含む各データ書込み要求において伝送される複数のボリューム及び情報を含むとき、書込みアドレス及びデータボリュームのIDの両方が合致すべき必要があることがわかる。

【0142】

ステップS318:見つかった書込みアドレスを含む情報に従って新しいデータ書込み要求を生成し、第2のキャッシュに新しいデータ書込み要求を書き込む。新しいデータ書込み要求は変更した番号を含む。

【0143】

例えば、第2のキャッシュから読み出される情報は、書込みアドレスA及び書き込むべきデータA(及び、セカンダリボリュームIDをも含み得る)を含み、新しいデータ書込み要求は、読み出される情報及び変更された番号(例えば、番号は13から14に変更される)に従って生成され得る。新しいデータ書込み要求が第2のキャッシュに書き込まれた後のキャッシュに格納されたマッピング関係が、表6に示される。

【0144】

【表6】

表 6

番号	ボリューム ID	書込みアドレス	書き込むべきデータ
12	セカンダリボリューム ID	書込みアドレス A	書き込むべきデータ A
12	セカンダリボリューム ID	書込みアドレス B	書き込むべきデータ D
12	セカンダリボリューム ID	書込みアドレス C	書き込むべきデータ C
13	セカンダリボリューム ID	書込みアドレス A	書き込むべきデータ E
14	セカンダリボリューム ID	書込みアドレス A	書き込むべきデータ A

【0145】

ホストが障害復旧アレイにデータ読出し要求を送信するとき、要求されたデータボリュームのIDがセカンダリボリュームIDであるとともに、書込みアドレスが書込みアドレスA

10

20

30

40

50

であり、障害復旧アレイは、セカンダリボリュームID及び書込みアドレスAを有し、最新の番号に対応する書き込むべきデータを第2のキャッシュから探索し、ホストに書き込むべきデータを送信する。本発明の実施形態では、番号14に対応する書き込むべきデータAが第2のキャッシュからホストに送信される。

【0146】

本発明の実施形態では、生産アレイが、データボリュームから関連情報を読み出す必要なく、キャッシュから障害復旧アレイに受信したデータ書込み要求において伝送される情報を直接送信し得、そのことがデータ複製の効率を上げるとともに、障害復旧アレイのデータ一貫性を保証する。

【0147】

従来技術では、データ複製は、スナップショットデータによって実施され、その場合、生産アレイがデータ書込み要求を実行する毎に、キャッシュ内にデータ書込み要求において伝送されるデータをまず配置し、アドレスに格納された古いデータをデータ書込み要求において伝送される書込みアドレスに従って読み出し、データボリューム内にデータを格納し、書込みアドレスにキャッシュ内のデータを書き込む必要がある。データ書込み要求の応答メッセージは、これら全ての動作が完了するまでに返されない。追加のスナップショットプロセスのステップのために、データ書込み要求の処理の遅延が長引く。しかしながら、本発明の実施形態では、スナップショット処理がデータ上で実行される必要がなく、従って、データ書込み要求が変更されるとはいえ、消費時間が短い。従って、従来技術と比べ、本発明の実施形態は、データ書込み要求を処理する遅延を減少する。

【0148】

図5を参照すると、図5は、本発明の実施形態によるストレージデバイス50の概略の構造図である。図5に示されるように、ストレージデバイス50は、受信モジュール501、読出し及び書込みモジュール502、カレント時間間隔番号マネージャ503、及び送信モジュール504を含む。

【0149】

受信モジュール501は、ホストによって送信される第1のデータ書込み要求を受信するように構成され、第1のデータ書込み要求は、書き込むべきデータ及びアドレス情報を伝送する。

【0150】

アドレス情報は、論理ブロックアドレス(LBA)を含み得る。
ストレージデバイス50が複数のデータボリュームを含むとき、アドレス情報は、ストレージデバイス50のデータボリュームのIDをさらに含む。

【0151】

読出し及び書込みモジュール502は、書き込むべきデータ及びアドレス情報に第1の番号を付与するとともに、キャッシュに情報を書き込むように構成され、第1の番号はカレント時間間隔番号であり、キャッシュから第1の番号に対応する書き込むべきデータ及びアドレス情報を読み出すように構成される。

【0152】

ストレージデバイス50は、カレント時間間隔番号マネージャ503を含み得、カレント時間間隔番号マネージャ503は、カレント時間間隔番号を格納し、カレント時間間隔番号は、0、1、2などの数字、又は、a、b、cなどの文字でよく、本明細書では限定されない。

【0153】

第1のデータ書込み要求が受信されるとき、第1の番号が、第1のデータ書込み要求において伝送される書き込むべきデータ及びアドレス情報に追加され、第1の番号は、カレント時間間隔番号によって割り当てられた値である。

【0154】

第1の番号が第1のデータ書込み要求において伝送される情報に付与された後、第1のデータ書込み要求において伝送される変更された情報がキャッシュに書き込まれ、その結果、書き込むべきデータ、アドレス情報、及び第1のデータ書込み要求において伝送される

10

20

30

40

50

第1の番号が、全てキャッシュに格納される。

【0155】

加えて、時間期間内に、別のデータ書込み要求がまた受信されることがあり、データ書込み要求において伝送される情報に第1の番号を付与するとともに、キャッシュに情報を書き込む必要がある。カレント時間間隔番号が変わる前に、第1の番号が全てのデータ書込み要求において伝送される情報に付与されることに留意すべきである。複製タスクがトリガされるとき、ストレージデバイス50は第1の番号に対応する書き込むべきデータ及びアドレス情報をキャッシュから読み出し得る。第1の番号に対応する書き込むべきデータ及びアドレス情報の複数の部分があることがわかる。

【0156】

複製タスクは、ストレージデバイス50が時間期間内にデータボリュームによって受信されるデータ書込み要求において伝送される情報を、障害復旧センタ内のストレージデバイスに送信し、カレント時間間隔番号と同じ番号が全てのデータ書込み要求において伝送される情報に付与される。複製タスクはタイマによってトリガされるか、又は手動によってトリガされ、本明細書では限定されない。複製の目的は、ストレージデバイス50によって受信されるデータ書込み要求において伝送される書き込むべきデータを、障害復旧センタ内のストレージデバイスに送信することであり、その結果、障害復旧センタ内のストレージデバイスは、ストレージデバイス50が故障したときに、ストレージデバイス50の働きを引き継ぐことができる。理解できるように、データ書込み要求において伝送されるアドレス情報(LBAなど)が、障害復旧センタ内のストレージデバイスに送信される必要があり、LBAは、障害復旧センタのストレージデバイスが書き込むべきデータを格納するアドレスを示すために使用される。障害復旧センタ内のストレージデバイスは、ストレージデバイス50と同様の物理構造を有する。従って、ストレージデバイス50に適用可能なLBAは、障害復旧センタ内のストレージデバイスにも適用可能である。

【0157】

本発明の実施形態では、複製タスクは、ストレージデバイス50のデータボリュームに特有である。ストレージデバイス50が複数のデータボリュームを含むとき、各データボリュームは1つの複製タスクに対応する。

【0158】

カレント時間間隔番号マネージャ503は、後続のデータ書込み要求において伝送される情報を識別するためにカレント時間間隔番号を変更するように構成される。

【0159】

複製タスクがトリガされるとき、カレント時間間隔番号マネージャ503はカレント時間間隔番号を変更する必要がある。後続のデータ書込み要求が受信されると、別の番号が、後続のデータ書込み要求において伝送される情報に付与される必要があり、別の番号は、変更されたカレント時間間隔番号によって割り当てられる値である。従って、キャッシュ内では、障害復旧センタ内のストレージデバイスに送信されるデータ書込み要求において伝送される情報は、ストレージデバイス50によって受信されるデータ書込み要求において伝送される情報と区別され得る。

【0160】

送信モジュール504は、障害復旧センタ内のストレージデバイスに書き込むべきデータ及びアドレス情報を送信するように構成される。

【0161】

ストレージデバイス50は、キャッシュから読み出した第1の番号に対応する書き込むべきデータ及びアドレス情報を、障害復旧センタ内の第2のストレージデバイスに送信する。

【0162】

具体的には、ストレージデバイス50は、障害復旧センタ内のストレージデバイスに、全ての読み出された書き込むべきデータ及びアドレス情報を直接送信し得る。又は、ストレージデバイス50は、障害復旧センタ内のストレージデバイスのデータボリュームのIDを取

10

20

30

40

50

得した後、各データ書込み要求において伝送される書き込むべきデータ及びアドレス情報に従って、新しいデータ書込み要求と、障害復旧センタ内のストレージデバイスのデータボリュームのIDとを生成し、障害復旧センタ内のストレージデバイスに新しいデータ書込み要求を送信する。

【0163】

本発明の実施形態では、ストレージデバイス50がホストによって送信されたデータ書込み要求を受信した後、データ書込み要求において伝送された情報が書き込むべきデータ及びアドレス情報を含み、ストレージデバイス50が第1の番号を書き込むべきデータ及びアドレス情報に付与し、キャッシュに情報を書込み、第1の番号はカレント時間間隔番号である。複製タスクがトリガされると、ストレージデバイス50が、キャッシュから第1の番号に対応する書き込むべきデータ及びアドレス情報を読み出し、障害復旧センタ内のストレージデバイスにその情報を送信する。加えて、複製タスクがトリガされると、ストレージデバイス50が、カレント時間間隔番号を変更し、その結果、ストレージデバイス50は、続いてデータ書込み要求を受信するときに、データ書込み要求において伝送される情報に、変更されたカレント時間間隔番号と同じ番号を付与する。それによって、キャッシュ内では、障害復旧センタ内のストレージデバイスに送信されたデータ書込み要求において伝送される情報は、ストレージデバイス50によって受信されるデータ書込み要求において伝送される情報と区別される。このことは、キャッシュから障害復旧センタ内のストレージデバイスへのデータ書込み要求において伝送される情報の直接送信を実施する。情報が、データボリュームから読み出す必要なく、キャッシュから直接送信されるので、データ複製時間が短く、そのことがデータ複製の効率を上げる。

【0164】

図6を参照すると、図6は、本発明の実施形態によるストレージデバイス60の概略の構成図である。

【0165】

図6に示されるように、ストレージデバイス60は、受信モジュール601、探索モジュール602、及び書込みモジュール604を含む。

【0166】

受信モジュール601は、ストレージデバイス50によって送信されたアドレス情報を受信するように構成される。

【0167】

具体的には、ストレージデバイス60は、ストレージデバイス50によって送信された書き込むべきデータ及びアドレス情報を受信し得、又はストレージデバイス60は、ストレージデバイス50によって送信されたデータ書込み要求を受信し得、データ書込み要求は書き込むべきデータ及びアドレス情報を含み、アドレス情報は、論理ブロックアドレス(LBA)であり得る。ストレージデバイス60が複数のデータボリュームを含むとき、アドレス情報は、ストレージデバイス60のデータボリュームのIDをさらに含み得る。1つ以上のアドレス情報があり得ることわかる。

【0168】

書き込むべきデータ及びアドレス情報を受信した後、ストレージデバイス60は、カレント時間間隔番号と同じ番号を書き込むべきデータ及びアドレス情報に付与し、その情報をキャッシュに書込み、その結果、カレント時間間隔番号と同じ番号、書き込むべきデータ、及びアドレス情報がキャッシュに格納される。

【0169】

ストレージデバイス60はまた、カレント時間間隔番号マネージャ603を含み得、カレント時間間隔番号マネージャ603は、カレント時間間隔番号を格納することに留意すべきである。カレント時間間隔番号は、0、1、2などの数字、又は、a、b、cなどの文字でよく、本明細書では限定されない。ここでのカレント時間間隔番号は、ストレージデバイス50内のカレント時間間隔番号と関係がなくともよい。

【0170】

探索モジュール602は、ストレージデバイス50が故障していることが決定すると、ストレージデバイス60がアドレス情報に従って第1の番号に対応する書き込むべきデータを取得するように構成され、第1の番号に対応するアドレス情報は、受信したアドレス情報と同じであり、第1の番号は、カレント時間間隔番号に先行する番号である。

【0171】

一般に、ストレージデバイス50及びストレージデバイス60の両方が、正常に動作する場合、ストレージデバイス60は、ストレージデバイス50によって送信された全てのデータ書き込み要求において伝送される情報を受信してよく、データ書き込み要求において伝送される情報に、カレント時間間隔番号と同じ番号を付与し、かつキャッシュにその情報を格納する。しかしながら、ストレージデバイス50が故障する場合、ストレージデバイス60は、ストレージデバイス50のカレント時間間隔番号に対応する書き込むべきデータの一部のみを受信することがある。この場合、ストレージデバイス60によって格納されるデータは、不正確である可能性があり、ストレージデバイス60がストレージデバイス50の働きを直接引き継ぐ場合、データの一貫性が保証できない。例えば、ホストは、この時点でストレージデバイス60にデータ読出し要求を送信する場合、アドレス情報(LBAなど)に格納された読出しデータを要求すること、ストレージデバイス60は、アドレス情報に対応する最新の番号を探索し、カレント時間間隔番号に対応する書き込むべきデータをホストに送信し、しかしながら、データは不正確である。従って、この場合、ストレージデバイス60のキャッシュ内のデータから、ストレージデバイス60のカレント時間間隔番号に先行する番号に対応するデータをリストアする必要がある。

10

20

【0172】

具体的に、ストレージデバイス50が故障することが、制御センタがストレージデバイス60に信号を送信する方法によって判断され得る。信号はストレージデバイス50が故障したことを示すために使用され、ストレージデバイス60は、ストレージデバイス50のホストサービスを引き継ぐ必要がある。

【0173】

一般に、複製タスクが完了するとき、制御センタは、ストレージデバイス50及びストレージデバイス60の両方に複製成功指示を送信し得る。ストレージデバイス60が指示を受信しない場合、現在の複製タスクが完了しないことを示す。複製タスクの完了は、ストレージデバイス50がカレント時間間隔番号に対応する全てのデータ書き込み要求で伝送される情報をストレージデバイス60に送信し、ストレージデバイス60がその情報の受信を終了したことを意味する。

30

【0174】

ストレージデバイス60は、ストレージデバイス50が故障していることが決定すると、現在の複製タスクが完了する場合に、ストレージデバイス60がストレージデバイス50の働きを直接引き継ぐことができ、データの一貫性が保証され得る。この状況は、本発明の実施形態の議論の範囲外である。

【0175】

しかしながら、現在の複製タスクが完了しない場合、ストレージデバイス60のキャッシュ内のデータから、ストレージデバイス60のカレント時間間隔番号に先行する番号に対応するデータをリストアする必要がある。

40

【0176】

具体的な複製方法は、受信したアドレス情報に従って、受信したアドレス情報と同じアドレス情報のために、カレント時間間隔番号に先行する番号に対応するアドレス情報を探索し、同じアドレス情報が見つからない場合、アドレス情報が見つかるまでさらに先行する番号に対応するアドレス情報を探索することを継続し、番号に対応する書き込むべきデータを取得する。

【0177】

書き込みモジュール604は、第1の番号に対応する書き込むべきデータ及びアドレス情報に第2の番号を付与し、キャッシュにその情報を書き込むように構成される。

50

【0178】

第2の番号は、カレント時間間隔番号を変更することによって取得された番号であり、この実施形態では、キャッシュ内に格納された最新の番号でもある。ホストがストレージデバイス60にデータ読み出し要求を送信するとき、アドレス情報(LBAなど)内に格納された読み出しデータを要求し、ストレージデバイス60は、アドレス情報に対応する最新の番号が第2の番号であることを探索することによって学習し、第2の番号に対応する書き込むべきデータをホストに送信する。それによって、データの一貫性が保証される。

【0179】

本発明の実施形態では、ストレージデバイス60は、ストレージデバイス50によって送信されたアドレス情報を受信し、ストレージデバイス50が故障するとき、カレント時間間隔番号に先行する番号に対応する書き込むべきデータを取得し、カレント時間間隔番号に先行する番号に対応する書き込むべきデータ及びアドレス情報に、第2の番号を付与し、キャッシュにその情報を格納する。それによって、データの一貫性が保証される。

10

【0180】

図7を参照すると、本発明の実施形態として、ストレージデバイス700が提供される。ストレージデバイス700は、従来技術で既知であるストレージデバイスを含み得る。本発明の実施形態は、ストレージデバイス700の特定の実装に限定されない。ストレージデバイス700は、プロセッサ710、通信インタフェース720、メモリ730、及び通信バス740を含む。

【0181】

プロセッサ710、通信インタフェース720、及びメモリ730は、通信バス740を使用することによって相互通信を実施する。

20

【0182】

通信インタフェース720は、例えば、ホスト又はスイッチと、ネットワーク要素と通信するように構成される。

【0183】

プロセッサ710は、プログラム732を実行するように構成される。
具体的には、プログラム732はプログラムコードを含み、プログラムコードはコンピュータ動作命令を含む。

【0184】

プロセッサ710は、CPU(central processing unit)、又はASIC(Application Specific Integrated Circuit)であり、本発明の実施形態を実装する1つ又は複数の組み込み回路に構成される。

30

【0185】

メモリ730は、プログラム732を格納するように構成される。
メモリ730は、高速RAM、不揮発性メモリ、例えば、少なくとも1つの磁気ディスクメモリを含み得る。

【0186】

プログラム732は、具体的に、ホストによって送信された第1のデータ書き込み要求を受信するように構成される受信モジュール501であって、第1のデータ書き込み要求は、書き込むべきデータ及びアドレス情報を伝送する、受信モジュールと、書き込むべきデータ及びアドレス情報に第1の番号を付与し、キャッシュにその情報を書き込む、読み出し及び書き込みモジュール502であって、第1の番号はカレント時間間隔番号であり、キャッシュから第1の番号に対応する書き込むべきデータ及びアドレス情報を読み出す、読み出し及び書き込みモジュールと、後続のデータ書き込み要求において伝送される情報に一致するように、カレント時間間隔番号を変更するように構成されるカレント時間間隔番号マネージャ503と、障害復旧センタ内のストレージデバイスに書き込むべきデータ及びアドレス情報を送信するように構成される送信モジュール504とを含み得る。

40

【0187】

プログラム732内のモジュールの具体的な実装のために、図5に説明される実施形態の対

50

応するモジュールが参照されてよく、本明細書にはさらに記載しない。

【0188】

図8に参照されるように、本発明の実施形態は、ストレージデバイス800を提供する。ストレージデバイス800は、従来技術で既知であるストレージデバイスを含み得る。本発明の実施形態は、ストレージデバイス800の特定の実装を限定しない。ストレージデバイス800は、プロセッサ810、通信インタフェース820、メモリ830、及び通信バス840を含む。

【0189】

プロセッサ810、通信インタフェース820、及びメモリ830は、通信バス840を使用することによって相互通信を実施する。

【0190】

通信インタフェース820は、例えば、ホスト又はスイッチである、ネットワーク要素と通信するように構成される。

【0191】

プロセッサ810は、プログラム832を実行するように構成される。

具体的には、プログラム832はプログラムコードを含み、プログラムコードはコンピュータ動作命令を含む。

【0192】

プロセッサ810は、CPU(central processing unit)、又はASIC(Application Specific Integrated Circuit)であり得、本発明の実施形態を実装する1つ又は複数の組み込み回路に構成され得る。

【0193】

メモリ830は、プログラム832を格納するように構成される。

メモリ830は、高速RAM、不揮発性メモリ、例えば、少なくとも1つの磁気ディスクメモリを含み得る。

【0194】

プログラム832は、具体的には、ストレージデバイス50によって送信されるアドレス情報を受信するように構成される受信モジュール601と、ストレージデバイス50が故障していることが決定すると、ストレージデバイス60が、アドレス情報に従って第1の番号に対応する書き込むべきデータを取得するように構成される探索モジュールであって、第1の番号に対応するアドレス情報は、受信したアドレス情報と同じであり、第1の番号がカレント時間間隔番号に先行する番号である、探索モジュール602と、第1の番号に対応する書き込むべきデータ及びアドレス情報に第2の番号を付与し、その情報をキャッシュに書き込むように構成される書込みモジュール604とを含み得る。

【0195】

プログラム832内のモジュールの具体的な実装のために、図6に説明される実施形態の対応するモジュールが参照されてよく、本明細書にはさらに記載しない。

【0196】

当業者に明確に理解できるように、好都合で明確な説明の目的で、上述のデバイス及びモジュールの詳細な動作プロセスに関して、上述の方法の実施形態における対応する処理が参照されてよく、本明細書では詳細を再度記載しない。

【0197】

少なくとも2つの障害復旧センタを含むシナリオに対する本発明のこの実施形態で提供される方法のアプリケーションが以下に記載される。

【0198】

図9に示されるように、ストレージシステムは、1つの生産センタ及び少なくとも2つの障害復旧センタを含む。

【0199】

生産センタは、生産ホスト、接続デバイス、及び生産アレイを含む。障害復旧センタのシステムのアーキテクチャは、障害復旧ホスト、接続デバイス、及び障害復旧アレイを含む、生産センタのアーキテクチャと同様である。生産センタ及び障害復旧センタは、IP(I

10

20

30

40

50

nternet Protocol)又はFC(Fiber Chanel)を介してデータ転送を実行し得る。生産センタ及び障害復旧センタの間に制御センタがあり得る。制御センタは、生産センタ側又は障害復旧センタ側に配置され得るか、生産センタ及び障害復旧センタの間にサードパーティデバイスに配置され得る。制御センタは、生産アレイが故障するとき、障害復旧アレイに信号を送信するように構成され、その結果、障害復旧アレイは生産アレイのホストサービスを引き継ぐ。

【0200】

生産ホスト及び障害復旧ホストは、サービス及びデスクトップなど従来技術で既知である任意のコンピュータであり得る。オペレーティングシステム及び他のアプリケーションプログラムがホスト内部にインストールされる。

10

【0201】

接続デバイスは、ファイバスイッチ又は他の現在利用可能なスイッチなど、ストレージデバイス及びホストの間に、従来技術で既知である任意のインタフェースを含み得る。

【0202】

生産アレイ及び障害復旧アレイは、1つ又は複数の相互接続したRAID(Redundant Array of Inexpensive Disks)、JBOD(Just a Bunch Of Disks)、及びテープライブラリなどの1つ又は複数の相互接続したディスクドライブなどの、従来技術で既知であるストレージデバイスであるか、又は、1つ又は複数の相互接続したディスクドライブがDASD(Direct Access ストレージデバイス)のものである、1つ又は複数のストレージユニットのテープストレージデバイスの両方であり得る。

20

【0203】

生産アレイのストレージ領域は、複数のデータボリュームを含み得る。データボリュームは、物理ストレージ領域をマッピングすることによって形成される論理ストレージ領域である。例えば、データボリュームは、LUN(Logic Unit Number)又はファイルシステムである。本発明のこの実施形態では、障害復旧アレイは生産アレイの構造と同様の構造を有する。

【0204】

生産センタ内のデータの安全性を保証するために、一般に、複数の障害復旧センタが配置される必要がある。生産センタ内のデータは複数の障害復旧センタに複製され、その結果、生産センタ及び1つの障害復旧センタの両方が障害に遭遇するとき、別の障害復旧センタがデータを格納し、データが失われない。本発明のこの実施形態では、生産アレイが障害復旧アレイに1つのデータボリュームに格納するデータを生産アレイが複製するタスクは、複製関係(又はペア)と呼ばれる。各複製関係は、1つのユニークな識別子(例えば、ID)に対応する。障害が発生する前に、生産アレイはホストによって送信されたデータ書き込み要求を継続的に受信するので、生産アレイは、障害復旧アレイに格納するデータを継続的に複製する必要がある。従って、1つの複製関係は、時間期間の一定の番号に分割されてよく、生産アレイが、各期間内で1つのデータボリュームによって受信されるデータ書き込み要求において伝送される情報を送信するタスクは、複製タスクと呼ばれる。

30

【0205】

生産アレイは、カレント時間間隔番号を格納するカレント時間間隔番号マネージャを含み得る。カレント時間間隔番号は、本明細書では限定しないが、0、1、2などの数字、又は、a、b、cなどの文字であり得る。カレント時間間隔番号は、各障害復旧アレイに適用可能であることに留意すべきである。複製タスクがトリガされる毎に、カレント時間間隔番号は変更される。

40

【0206】

例として、カレント時間間隔番号が1であることが使用される。生産アレイがデータ書き込み要求を受信するとき、数値1の番号がデータ書き込み要求において伝送される書き込むべきデータ及びアドレス情報に付与され、データ書き込み要求において伝送される書き込むべきデータ及びアドレス情報、及び数値1の番号は、全てキャッシュに書き込まれる。

【0207】

50

第1の障害復旧アレイに対応する複製タスクがトリガされるとき、生産アレイは数値1から数値2にカレント時間間隔番号を変更し、その結果、数値2の番号が、後続の受信したデータ書込み要求において伝送される書き込むべきデータ及びアドレス情報に付与される。その後、数値1の番号に対応する書き込むべきデータ及びアドレス情報が、第1の障害復旧アレイに送信される。

【0208】

第2の障害復旧アレイに対応する複製タスクがトリガされるとき、生産アレイは、数値2から数値3にカレント時間間隔番号を変更し、その結果、数値3の番号が、後続の受信したデータ書込み要求において伝送される書き込むべきデータ及びアドレス情報に付与される。

10

【0209】

第3の障害復旧アレイに対応する複製タスクがトリガされるとき、生産アレイは、数値3から数値4にカレント時間間隔番号を変更し、その結果、数値4の番号が後続の受信したデータ書込み要求において伝送される書き込むべきデータ及びアドレス情報に付与される。

【0210】

しかしながら、第1の障害復旧アレイに対応する複製タスクが別の時間のためにトリガされるとき、複製アレイは、数値4から数値5にカレント時間間隔番号を変更し、その結果、数値5の番号が、後続の受信したデータ書込み要求において伝送される書き込むべきデータ及びアドレス情報に付与される。生産アレイはカレント時間間隔番号のみを記録すると仮定する。その後、この時点で、それに応じて、数値4の番号に対応する書き込むべきデータ及びアドレス情報が第1の障害復旧センタに送信される。結果的に、数値2の番号に対応する書き込むべきデータ及びアドレス情報と、数値3の番号に対応する書き込むべきデータ及びアドレス情報とは省略され、第1の障害復旧アレイに格納されたデータは、生産アレイ内のデータとの一貫性がなくなる。同様に、第2の障害復旧アレイ及び第3の障害復旧アレイは、受信した書き込むべきデータ及びアドレス情報が不完全であるという問題にも直面する。

20

【0211】

図10を参照すると、図10では、その問題を解決するための、本発明によって提案されるデータ複製方法の実施形態を示す。方法は、少なくとも2つの障害復旧アレイに対応する生産アレイに適用される。本発明のこの実施形態では、記載を容易にするために、生産アレイが、第1のストレージデバイスとして参照され、少なくとも2つの障害復旧アレイの1つが、第2のストレージデバイスとして参照される。以降のステップの具体的な実行のために図2から図4に示される実施形態が参照され得る。

30

【0212】

方法は、以下のステップを含む。

ステップ41：現在の複製タスクがトリガされるとき、第1のストレージデバイスは、カレント時間間隔番号を読み出す。複製タスクはタイマ又は手動によってトリガされてよく、又は他のトリガ方法でトリガされてよく、本明細書では限定されない。複製タスクがトリガされるとき、第1のストレージデバイスがカレント時間間隔番号マネージャからカレント時間間隔番号を読み出し得る。複製タスクがトリガされるとき、第1のストレージデバイスはトリガの瞬間にカレント時間間隔番号を変更し、本明細書では、第1のストレージデバイスによって読み出されたカレント時間間隔番号は、変更されたカレント時間間隔番号である。記載を容易にするために、以降のステップでは、変更前のカレント時間間隔番号を履歴時間間隔番号と呼ぶ。

40

【0213】

ステップ42：第2の番号を読み出す。第2の番号は、現在の複製タスクに相互に関連する最新の完了した複製タスクに対応する番号である。

【0214】

本発明のこの実施形態では、最新の完了した複製タスクが現在の複製タスクに相互に関連するということは、現在の複製タスク及び最新の完了した複製タスクが同一の複製関係

50

に属することを意味し、先行する記載から、各複製関係がユニークなIDを有することがわかる。複製タスクがトリガされるとき、第1のストレージデバイスはIDを受信し、IDに応じた第2の番号を読み出し得る。

【0215】

具体的には、現在の複製タスクがタイマによってトリガされるとき、IDはタイマ内で保有され得、現在の複製タスクが手動でトリガされるとき、第1のストレージデバイスは、信号の形態又は他の形態でIDを受信してもよい。

【0216】

本発明のこの実施形態では、複製タスクが完了する毎に、完了した複製タスクに対応する番号が記録される。限定ではないが、選択的に、最新の完了した複製タスクに対応する番号だけが格納されてよく、次の複製タスクが完了するときに番号が更新され、又は全ての完了した複製タスクに対応する番号が格納される。

10

【0217】

最新の完了した複製タスクに対応する番号が各複製関係に記録されることが理解できる。

【0218】

ステップ43：カレント時間間隔番号及び第2の番号に従って、第1の番号を決定する。第1の番号は、複製タスクがトリガされるときのカレント時間間隔番号より前の番号であり、第1の番号は第2の番号の後の番号である。

【0219】

例えば、カレント時間間隔番号は数値5であり、第2の番号は数値2であり、間隔(2,5)内の任意の番号が、第1の番号として決定され得る。区間は、数値2及び数値5を除く、開区間であることに留意すべきである。

20

【0220】

ステップ44：キャッシュに格納されるとともに第1の番号に対応する複製すべきデータ及び複製すべきデータのアドレス情報を、第2のストレージデバイスに複製する。

【0221】

具体的には、複製すべきデータと、第1の番号に対応する複製すべきデータのアドレス情報とがキャッシュから読み出され、複製すべきデータ及び複製すべきデータのアドレス情報は、第2のストレージデバイスに送信される。

30

【0222】

選択的に、第1のストレージデバイスは、複製すべきデータ及び複製すべきデータのアドレス情報を第2のストレージデバイスに直接送信し、又は複製すべきデータ及び複製すべきデータのアドレス情報に従ってデータ書込み要求を生成し得、データ書込み要求を第2のストレージデバイスに送信する。

【0223】

選択的に、現在の複製タスクがトリガされるとき、かつ最新の完了した複製タスクに対応する番号の後で、複数の番号は全て、カレント時間間隔番号の前である条件を満足し、同じアドレス情報に対応する複数の番号は、アドレス情報及び最新の番号に対応する複製すべきデータのみが、第2のストレージデバイスに送信され得る。最新の番号は、最も直前に生成された番号である。例えば、カレント時間間隔番号が各時間に1を付与することによって変更され、最新の番号は最も大きい数値をもつ番号であると仮定する。

40

【0224】

本発明のこの実施形態では、現在の複製タスクがトリガされるとき、第1のストレージデバイスは、カレント時間間隔番号及び第2の番号に従って第1の番号を決定し、第2の番号が、現在の複製タスクの前の最新の完了した複製タスクに対応する番号であり、第1の番号は、現在の複製タスクがトリガされるときのカレント時間間隔番号の前かつ第2の番号の後の番号であり、第1のストレージデバイスは、キャッシュに格納されるとともに第1の番号に対応する複製すべきデータ及び複製すべきデータのアドレス情報を第2のストレージデバイスに複製する。第2の番号とカレント時間間隔番号の間の全ての番号が、第1の

50

番号として決定され得るので、番号が第1の番号として決定されない限り、複製すべきデータと番号に対応する複製すべきデータのアドレス情報とが、第2のストレージデバイスに複製され得る。従って、別の障害復旧アレイに対応する複製タスクがトリガされる時、カレント時間間隔番号が変更される場合でさえ、第1のストレージデバイスが第2の番号に従って、複製すべきデータと第2のストレージデバイスに複製されない複製すべきデータのアドレス情報とを見つけ、それらを第2のストレージデバイスに複製し得る。そのことによって、複製の完全性が保証される。

【0225】

選択的に、上述の実施形態では、現在の複製タスクがトリガされる前に、方法は、第1のデータ書き込み要求を受信するステップであって、第1のデータ書き込み要求は、複製すべきデータ及び複製すべきデータのアドレス情報を含む、ステップと、第1の番号を複製すべきデータ及び複製すべきデータのアドレス情報に付与するステップと、キャッシュにそれらを書き込むステップとをさらに含み、第1の番号は、履歴時間間隔番号である。

10

【0226】

具体的には、履歴時間間隔番号は、第1のデータ書き込み要求が受信される瞬間に対応するカレント時間間隔番号を参照する。図2から図4に示される実施形態から、複製タスクがトリガされる時、履歴時間間隔番号はカレント時間間隔番号に変更されることが理解できる。

【0227】

選択的に、方法は、第2のデータ書き込み要求を受信するステップであって、第2のデータ書き込み要求は、ターゲットデータ及びターゲットデータのアドレス情報を含む、ステップと、第3の番号をターゲットデータ及びターゲットデータのアドレス情報に付与するステップと、第3の番号が第1の番号と同じであるか否かを判断するステップと、第3の番号が第1の番号と同じであるとき、ターゲットデータのアドレス情報が、複製されるべきデータのアドレス情報と同じか否かを判断するステップと、ターゲットデータのアドレス情報が、複製されるべきデータのアドレス情報と同じとき、キャッシュに格納された複製すべきデータをターゲットデータに置き換えるステップと、ターゲットデータのアドレス情報が複製されるべきデータのアドレス情報と同じでない場合、第3の番号が付与された後に、ターゲットデータ及びターゲットデータのアドレス情報をキャッシュに書き込むステップとをさらに含む。

20

30

【0228】

第3の番号が第1の番号と同じでない場合、第3の番号が付与された後に、ターゲットデータ及びターゲットデータのアドレス情報がキャッシュに書き込まれる。

【0229】

図11を参照すると、図11は、本発明の実施形態によるストレージデバイスの概略的な構造図である。図11に示されるように、ストレージデバイスは、読出し及び書き込みモジュール52、決定モジュール53、及び複製モジュール54を含む。

【0230】

読出し及び書き込みモジュール52は、現在の複製タスクがトリガされる時、カレント時間間隔番号を読み出すとともに、第2の番号を読み出すように構成され、第2の番号は、現在の複製タスクに相互に関連する最新の完了した複製タスクに対応する番号である。

40

【0231】

決定モジュール53は、カレント時間間隔番号及び第2の番号に従って、第1の番号を決定するように構成され、現在の複製タスクがトリガされる時、第1の番号は、カレント時間間隔番号の前の番号であり、第1の番号は第2の番号の後の番号である。

【0232】

複製モジュール54は、キャッシュ内に格納されるとともに第1の番号に対応する複製されるべきデータ及び複製されるべきデータのアドレス情報を、第2のストレージデバイスに複製するように構成される。

【0233】

50

本発明のこの実施形態では、現在の複製タスクがトリガされるとき、第1のストレージデバイスは、カレント時間間隔番号及び第2の番号に従って第1の番号を決定し、第2の番号は、現在の複製タスクの前の最新の完了した複製タスクに対応する番号であり、第1の番号は、現在の複製タスクがトリガされるときのカレント時間間隔番号の前かつ第2の番号の後の番号であり、複製されたデータ及びキャッシュに格納されかつ第1の番号に対応する複製されたデータのアドレス情報を第2のストレージデバイスに複製する。第2の番号とカレント時間間隔番号との間の全ての番号が、第1の番号として決定され得、番号が第1の番号として決定される限り、複製されるべきデータとその番号に対応する複製されるべきデータのアドレス情報とが、第2のストレージデバイスに複製され得る。従って、別の障害復旧アレイに対応する複製タスクがトリガされるときにカレント時間間隔番号が変更される場合でさえ、第1のストレージデバイスが、第2の番号に従って、複製されるべきデータと第2のストレージデバイスに複製されない複製されるべきデータのアドレス情報とを見つけ得、それらを第2のストレージデバイスに複製し得、それによって、複製の完全性が保証され得る。

10

【0234】

選択的に、ストレージデバイスが、第2の番号を記録するように構成される記録モジュール55を含み得る。

【0235】

選択的に、最新の完了した複製タスクが現在の複製タスクに相互に関連することは、現在の複製タスク及び最新の完了した複製タスクが、同じ複製関係に属していることを意味し、ストレージデバイスがさらに受信モジュール51を含む。

20

【0236】

受信モジュール51は、複製関係に対応する識別子を受信するように構成される。

読み出し及び書き込みモジュールは、具体的には、識別子に従って現在の複製タスクに対応する第2の番号を読み出すように構成される。

【0237】

選択的に、受信モジュール51は、現在の複製タスクがトリガされる前に、第1のデータ書き込み要求を受信するようにさらに構成され、第1のデータ書き込み要求は、複製されるべきデータ及び複製されるべきデータのアドレス情報を含み、読み出し及び書き込みモジュール52は、複製されるべきデータ及び複製されるべきデータのアドレス情報に第1の番号を付与し、それらをキャッシュに書き込むようにさらに構成され、第1の番号は、履歴時間間隔番号である。カレント時間間隔番号は、履歴時間間隔番号を変更することによって形成される。

30

【0238】

選択的に、複製モジュール54は、具体的には、アドレス情報が、複数の番号に対応するとき、アドレス情報に対応する番号のうちの最新の番号を第1の番号として決定し、キャッシュに格納されるとともに第1の番号に対応する複製されるべきデータ及び複製されるべきデータのアドレス情報を、第2のストレージデバイスに複製する。

【0239】

本発明のこの実施形態によって提供されるストレージデバイスは、上述の実施形態で説明されたデータ複製方法を実行するために使用される。モジュールの機能の詳細な記述のための方法の実施形態の記載が参照され、これ以上の詳細は本明細書には記載されない。

40

【0240】

図12に示されるように、図12では、本発明の実施形態によって提供されるストレージデバイスを示し、ストレージデバイスは、プロセッサ101、メモリ102、システムバス(ショートのためのバス)105、及び通信インタフェース103を含む。プロセッサ101、メモリ102、及び通信インタフェース103が接続され、システムバス105を使用することによって相互通信を実施する。

【0241】

プロセッサ101は、単一コア又は複数コアの中央処理ユニット、又は特定の集積回路で

50

あり得、本発明の実施形態を実施する1つ又は複数の集積回路に構成され得る。

【0242】

メモリ102は、高速RAMメモリ、又は不揮発性メモリ、例えば、少なくとも1つの磁気ディスクメモリであり得る。

【0243】

通信インタフェース103は、ストレージデバイスと通信するように構成される。

メモリ102は、コンピュータ実行可能命令1021を格納するように構成される。具体的には、コンピュータ実行可能命令1021は、プログラムコードを含み得る。

【0244】

コンピュータの実行時、プロセッサ101はコンピュータ実行可能命令1021を実行し、図10に説明される方法手順を実行することができる。 10

【0245】

本願で提供されるいくつかの実施形態では、開示されたデバイス及び方法は、他の方法で実施され得ることが理解されるべきである。例えば、記載された装置の実施形態は、単なる例示である。例えば、モジュールの分割は、単に、論理機能の分割であり、実際の実装では他の分割が可能である。例えば、複数のモジュール又は構成要素が、別のデバイスに組み合わされるか統合され得、又はいくつかの特徴が、無視されるか実行されなくてもよい。加えて、表示された又は議論された、相互結合、直接結合、又は通信接続は、いくつかのインタフェースを使用して実装され得る。装置間又はモジュール間の間接結合又は通信接続は、電氣的、機械的、又は他の形態で実装され得る。 20

【0246】

別々の部品で記載されたモジュールは、物理的に別々であってもなくてもよく、モジュールとして表示された部分は、物理的なサブモジュールであってもなくてもよく、1つの位置に位置してもよく、複数のネットワークサブモジュール上に分散されてもよい。モジュールのいくつか又は全ては、実際の必要性に従って実施形態の解決策を目的として達成するために選択され得る。

【0247】

加えて、本発明の実施形態の機能モジュールが、処理モジュールに統合されてもよく、又はモジュールの各々が物理的に単独で存在してもよく、2つ又はそれ以上のモジュールが1つのモジュールに統合されてもよい。 30

【0248】

当業者は、上述の実施形態のステップの全て又は一部が、ハードウェアによって実装されるか、又は関連ハードウェアに命令するプログラムによって実装され得ることを理解する。プログラムは、コンピュータ可読記録媒体に格納され得る。記録媒体は、読出し専用メモリ、磁気ディスク、又は光学ディスクであり得る。

【0249】

最後に、上述の実施形態は、単に、本発明の技術的解決策を記載するためのものであり、本発明を限定するものではないことに留意すべきである。本発明は、上述の実施形態を参照して詳細に記載されるが、そのような変更又は置き換えが、対応する技術的解決策の本質を本発明の実施形態の技術的解決策の範囲から逸脱しない限りで、当業者は、上述の実施形態に記載される技術的解決策に変更を加え、又は、その技術的特徴のいくつか又は全てに均等な置き換えが可能であることを理解すべきである。 40

【符号の説明】

【0250】

50 ストレージデバイス

501 受信モジュール

502 読み込み及び書き込みモジュール

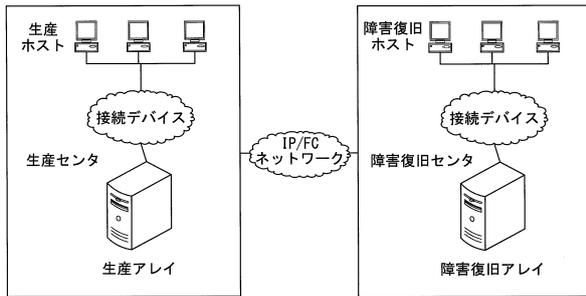
503 カレント時間間隔番号マネージャ

504 送信モジュール

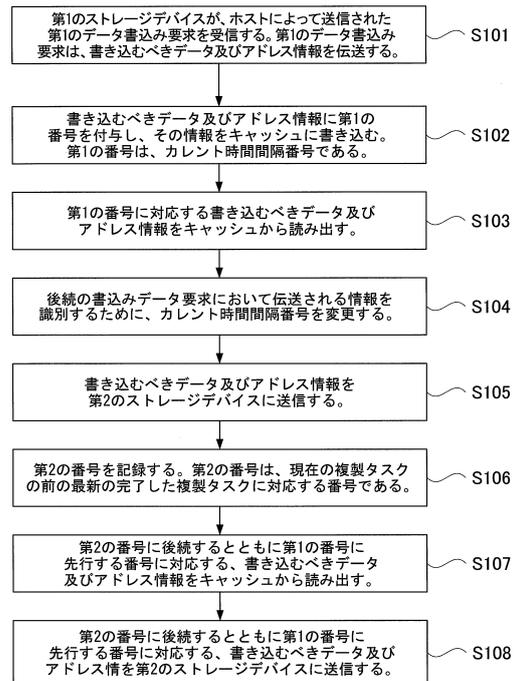
60 ストレージデバイス 50

- 601 受信モジュール
- 602 探索モジュール
- 603 カレント時間間隔番号マネージャ
- 604 書込みモジュール
- 700 ストレージデバイス
- 710 プロセッサ
- 720 通信インタフェース
- 730 メモリ
- 732 プログラム
- 740 通信バス
- 800 ストレージデバイス
- 810 プロセッサ
- 820 通信インタフェース
- 830 メモリ
- 832 プログラム
- 840 通信バス

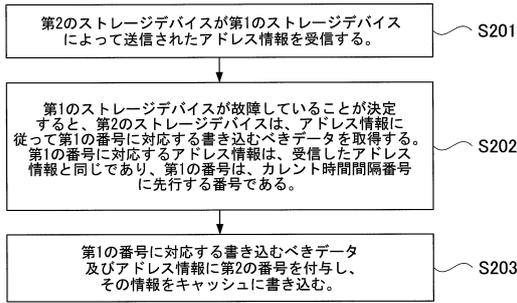
【図1】



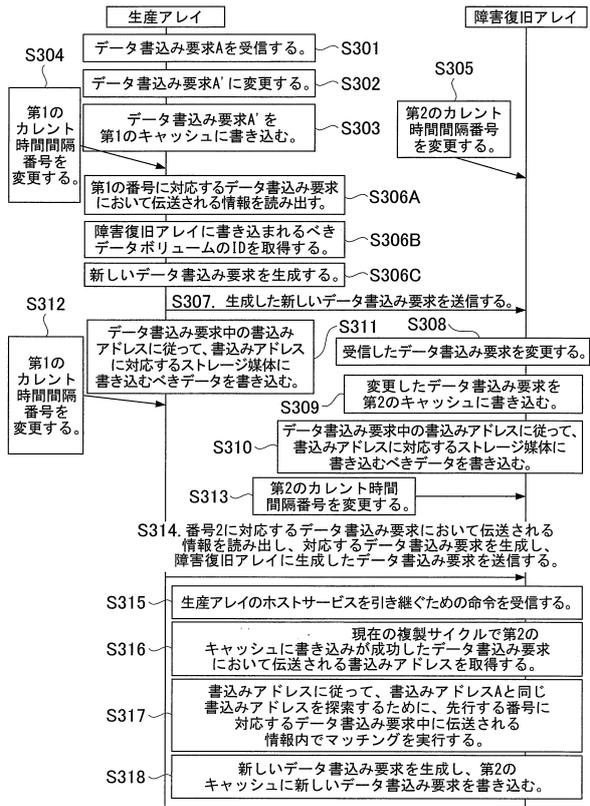
【図2】



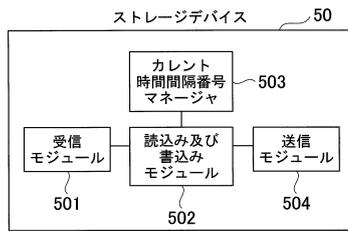
【図3】



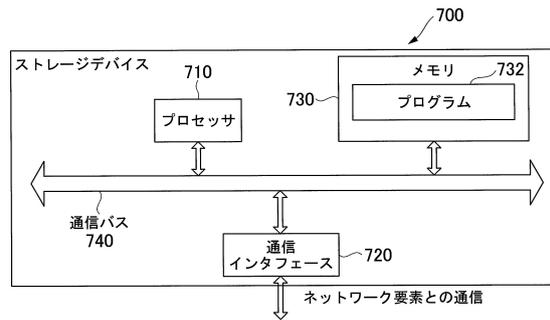
【図4】



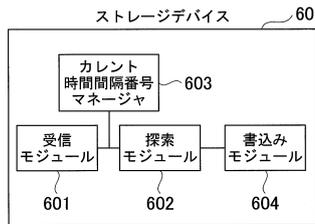
【図5】



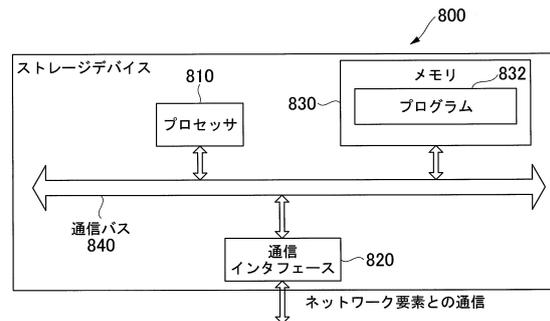
【図7】



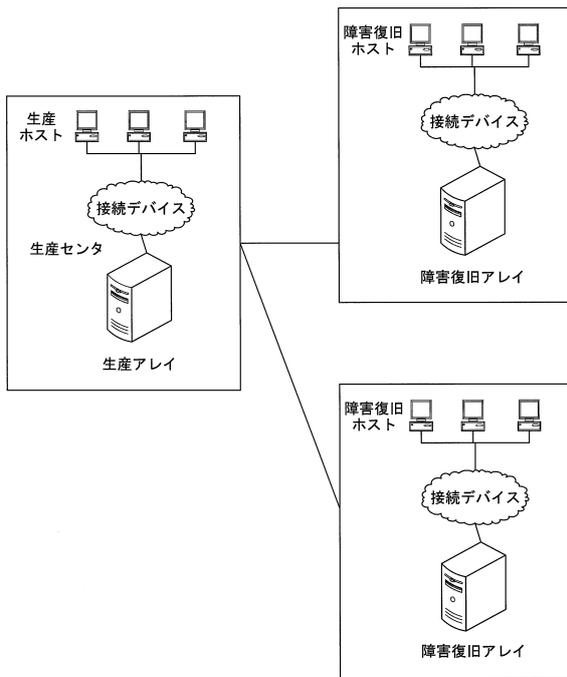
【図6】



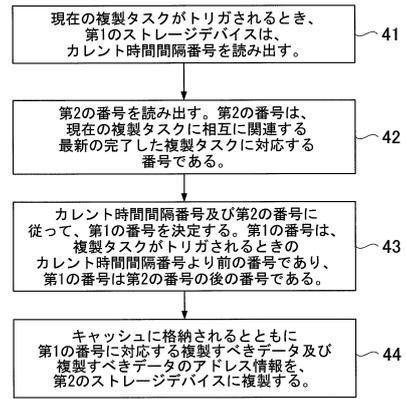
【図8】



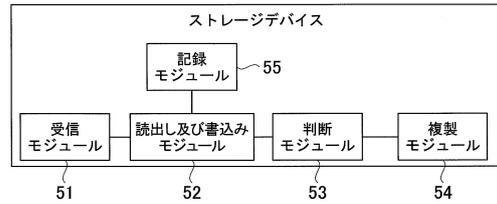
【図9】



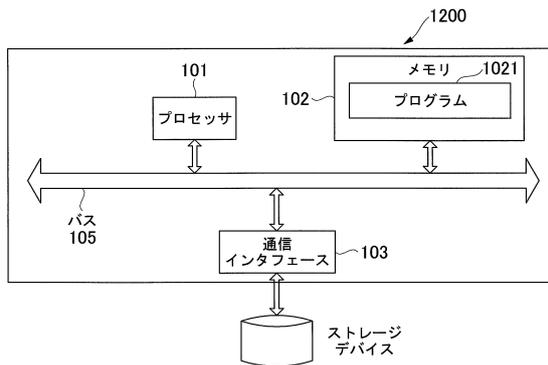
【図10】



【図11】



【図12】



フロントページの続き

- (72)発明者 欧陽 戟
中華人民共和国518129広東省深セン市龍岡区坂田華為本社ビル
- (72)発明者 鄒 冶
中華人民共和国518129広東省深セン市龍岡区坂田華為本社ビル

合議体

審判長 千葉 輝久

審判官 山澤 宏

審判官 稲葉 和生

- (56)参考文献 特開2012-003621(JP,A)
特開2007-066154(JP,A)
特開平07-244597(JP,A)
高橋 直也 Naoya TAKAHASHI, 「災害対策用ディスク間遠隔コピー方式の改良」, 電子情報通信学会論文誌 (J87-D-I) 第6号, 日本, 社団法人電子情報通信学会, 2004年6月1日, 第J87-D-I巻, P.675~692

- (58)調査した分野(Int.Cl., DB名)

G06F 3/06-3/08

G06F12/00

G06F13/00-13/14