

ストレージ横断リモートコピーにおけるライトデータ順序保証方式 Mechanisms Maintaining Write Data Consistency across Multiple Storages for Storage Based Remote Replication

牧晋広† 平岩友里† 渡辺和彦‡ 宮田和久‡
Nobuhiro Maki Yuri Hiraiwa Kazuhiko Watanabe Katsuhisa Miyata

1. まえがき

近年、企業における情報システムは大規模化し、データそのものの価値も高まる一方である。ところが、テロ、災害等によるシステム停止やデータ損失の発生を避けることは難しく、一度発生すると企業の損失は甚大である。そこで現在では、これらのシステム停止やデータ損失を回避するため、ディザスタリカバリが用いられる。ディザスタリカバリとは災害等により発生した情報システムの障害を復旧させる概念である[1]。

ディザスタリカバリを支える技術にストレージベースリモートコピーがある。ストレージベースリモートコピーとは、業務が実施される拠点（プライマリサイトとよぶ）のホスト計算機が生成したデータを、ストレージ装置が受信すると、当該ストレージ装置が遠隔拠点（セカンダリサイトとよぶ）のストレージ装置にコピーする方式である。セカンダリサイトのストレージ装置に業務データをコピーすることで、災害時もセカンダリサイトからの業務回復が可能となる[2]。

ストレージベースリモートコピーは、ストレージ装置単体によりデータのコピー処理を行うため、ホスト計算機にデータ転送の負荷を与えないメリットがある。その一方で、ストレージ装置単体でデータのコピー処理を行うため、複数のストレージ装置が連携してデータのコピー処理を行うことが難しい。結果として、データの処理性能向上のために、複数のストレージ装置にデータを分散配置する大規模アプリケーション環境への適用が難しくなる。

本報告では、複数のストレージ装置に渡ってデータの書き込み順序を保証してリモートコピーが可能となるストレージ横断リモートコピー制御方式に関し、特にリモートコピー開始直後における順序保証確立有無の判定方法を説明する。

2. 課題と目的

ストレージベースリモートコピーの制御の1つに、書き込みデータの順序保証制御がある。これは、ホスト計算機がストレージ装置内の複数論理ボリュームに発行した書き込みデータの書き込み順序を、セカンダリサイトのストレージ装置内の該当論理ボリュームにおいても保証する制御で、セカンダリサイトからの業務回復を可能にするためには必要な制御である。なぜならば、アプリケーションによっては論理ボリューム間に強い依存関係があり、その依存

関係が崩れると、場合によってはアプリケーションの起動が不可能になるためである。

また、リモートコピー開始直後は、書き込みデータの順序保証制御を可能にする処理を開始するが、当該処理が完了するまで、書き込みデータの順序は保証できない。そのため、アプリケーションの確実な回復を保証するためには、開始直後から書き込みデータの順序保証制御が確立するまでを監視する必要がある。

基本的にストレージベースリモートコピーは上記の制御をストレージ装置レベルで実装している。ストレージ装置レベルで実装することにより、ホスト計算機の負荷を削減できるメリットがある。その反面、複数ストレージ装置間で順序保証制御をする場合、リモートコピー開始直後は単独で動作する書き込みデータの順序保証制御を、リモートコピー対象の全ストレージ装置に渡って保証可能かを判定する必要がある。

本報告では、特に上記を解決する判定方法を示す。

3. ストレージ横断リモートコピー制御方式

3.1 方式概要

ストレージ横断リモートコピー制御方式では、ホスト計算機がデータに付与した時刻情報をリモートコピーの付加情報としてセカンダリサイトに転送し、セカンダリサイトでは全ストレージ装置のデータ書き込み順序の監視が可能な代表ストレージ装置を用意し、上記時刻情報を基に該当データをコピー先論理ボリュームに書き込む。

図1に本方式の構成を示す。プライマリサイト、セカンダリサイトには共に、複数のストレージ装置（図では各3台）がホスト計算機に接続されている。また、プライマリサイトのストレージ装置はセカンダリサイトのストレージ装置とデータネットワークを介して接続される。セカンダリサイトのストレージ装置のうち1つは代表ストレージ装

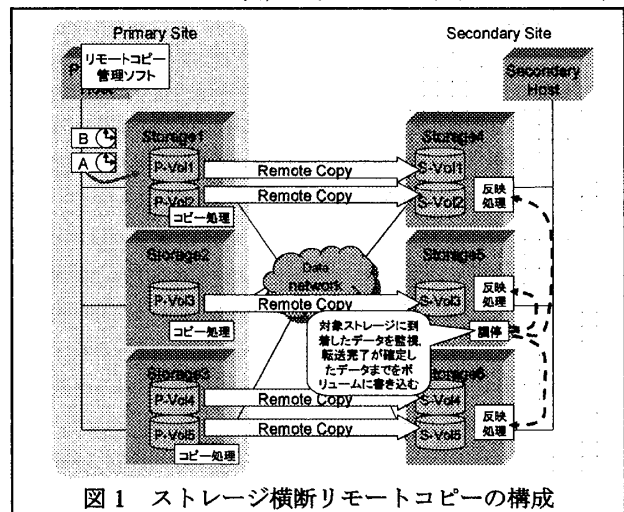


図1 ストレージ横断リモートコピーの構成

† (株)日立製作所 システム開発研究所
Systems Development Laboratory, Hitachi Ltd.
‡ (株)日立製作所 ソフトウェア事業部
Software Division, Hitachi Ltd.

置(図では Storage5)である。

プライマリサイトのホスト計算機には本方式のリモートコピーを管理するリモートコピー管理ソフトが動作する。リモートコピー管理にはリモートコピーを開始、一時停止などの制御と、リモートコピーの処理状況や障害を検出するためにその状態をストレージ装置から取得する監視がある。

プライマリサイトのストレージ装置は、ホスト計算機から受信したデータをコピーし、セカンダリサイトに転送するコピー処理部を持つ。

セカンダリサイトのストレージ装置は、転送されたデータを一時保存し、代表ストレージ装置の指示に応じて、該当論理ボリュームに書き込む反映処理部をもつ。代表ストレージ装置は、反映処理部に加え、セカンダリサイトの全ストレージ装置に対して指示が可能な調停処理部を持つ。

ここで、上記課題を解決するため、リモートコピー開始直後、コピー元とコピー先のストレージ装置で単独に動作するリモートコピー処理を連携する代表ストレージ装置の結合処理の進捗と、各ストレージ装置が管理するリモートコピーの状態の両方を考慮して、リモートコピー管理ソフトで書き込みデータの順序保証制御の確立を判定する。

上記処理により、リモートコピーの順序保証制御の確立判定が可能になり、課題を解決する。

3.2 開始直後のリモートコピー処理と書き込みデータ順序保証の判定方法

リモートコピー開始直後の処理について、図2を用いて説明する。本方式を含め一般的に、リモートコピー制御では、開始直後にコピー元の論理ボリュームとコピー先の論理ボリュームの内容一致のためのコピー処理を実施する。この処理は初期コピーとよばれる。初期コピー期間中はホスト計算機から発行されたデータではなく、論理ボリューム内のデータを転送することになるため、書き込み順序の保証が不可能になる。この期間のコピーの状態を初期コピー状態とよび、プライマリサイトの各ストレージ装置でコピー状態の進捗を管理し、リモートコピー状態として保持する。

一方、本方式では初期コピー中に、初期コピー処理に平行して、結合処理を実施する。具体的には、代表ストレージ装置がセカンダリサイトの対象ストレージ装置に対し、代表ストレージ装置と対象ストレージ装置間の通信可否の確認し、初期コピー完了後に転送されたデータを代表ストレージ装置からの指示無く、コピー先論理ボリュームに書き込まないように指示する、といった処理を実施する。代表ストレージ装置はこの結合処理の進捗を管理し、一覧情報として保持する。

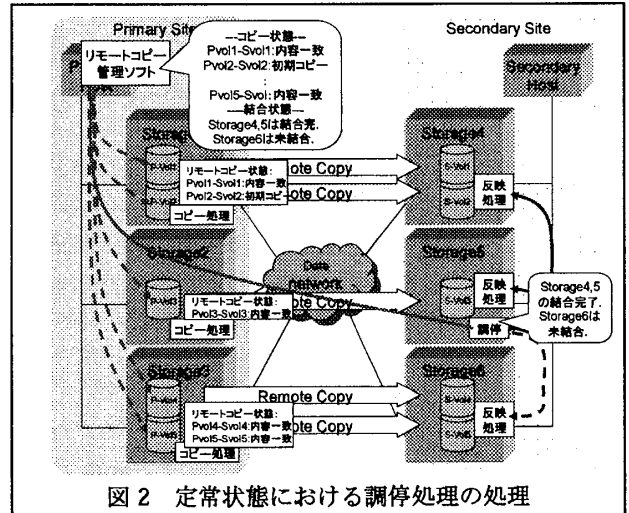


図2 定常状態における調停処理の処理

また、リモートコピー管理ソフトはリモートコピーの監視のため、プライマリサイトのストレージ装置からコピー状態を取得し、代表ストレージ装置から結合処理の進捗を取得する。上記の情報取得後、リモートコピー管理ソフトは表1を用いて書き込みデータの順序保証について判定する。すなわち、リモートコピー状態が内容一致であったとしても、結合の状態が未結合の場合、リモートコピーの状態を初期コピー状態と判定し、書き込みデータの順序保証ができない状態とする。

上記の判定処理により、初期コピーが完了しているが、複数のストレージ装置にわたって書き込みデータの順序が保証されていない状態を、書き込みデータの順序確立として誤判定しないようにできる。

4. むすび

本報告では、複数のストレージ装置に渡ってデータの書き込み順序を保証してリモートコピーが可能なストレージ横断リモートコピー制御方式で、特にリモートコピー開始直後における順序保証確立有無の判定方法を説明した。

本判定方法は、ストレージ装置が持つリモートコピー状態と代表ストレージ装置が持つ結合処理の進捗をリモートコピー管理ソフトで総括して、データの書き込み順序の判定を実施する。そのため、ストレージ装置のみでは誤判定するリスクを抑止することが可能となる。

本ストレージ横断リモートコピー制御方式は2サイト間でのリモートコピー構成が基本である。しかし、近年ではより高いデータ回復性を可能にする複数サイト構成のリモートコピーが注目されているため、本方式の複数サイト構成への適用は今後の課題である。

参考文献

[1] Rudolph, C. G, "Business continuation planning/disaster recovery", Communications Magazine, IEEE, Vol.28, Issue6, PP. 25-28, June 1990
 [2] Advanced functions for storage subsystems: Supporting continuous availability, IBM SYSTEM JOURNAL, Vol.42, No.2, P.268, 2003

表1 リモートコピー管理ソフトによる順序保証判定基準

		結合処理	
		未結合	結合
リモートコピー状態	初期コピー	初期コピー	初期コピー
	内容一致	初期コピー	内容一致