

3 拠点カスケードリモートコピー構成における 主ホスト障害時のデータ可用性維持方式

岡田 渡[†] 牧 晋広[†] 細内 昌明[†] 平岩友里[†] 宮田 和彦[‡] 佐藤 雅英[‡]

(株)日立製作所 システム開発研究所[†]

(株)日立製作所 ソフトウェア事業部[‡]

1. はじめに

IT 普及により企業情報システムの重要性は高まる一方である。ところが、テロ、災害等によるシステム停止やデータ損失は発生しないわけではない。一旦システム停止や業務データの損失が発生すると、企業収益に影響を与えるばかりでなく、企業の信頼性を失いかねないため、上記災害への対策は必要不可欠である[1]。

システム停止やデータ損失のリスクを回避するため、ディザスタリカバリ技術が用いられる。近年ではデータを保管するストレージシステム（以下単にストレージ）におけるディザスタリカバリ技術が注目されている。その中でも特に、データ損失のリスクを極小化できる3拠点カスケードリモートコピー構成（以下単に Cascade 構成とよぶ）が注目されている。

Cascade 構成は短距離伝送だがデータ損失の無い同期コピーで業務データをコピーし、さらに当該データをデータ損失はあるが長距離伝送が可能な非同期コピーでコピーする。これにより、業務ホスト近辺の小規模災害はデータ損失をゼロに、広域災害時もデータ損失を最小にする。ところで上記のような災害時、3拠点カスケードリモートコピーは最遠地から業務回復を行う。業務回復後は拠点間の距離関係が異なるため障害前と同じコピー構成を使用することができず、データ可用性が低下する。

本報告では障害前後で3拠点カスケードリモートコピーのデータ可用性を維持する方式を提案する。

2. Cascade 構成とその課題

2.1 Cascade 構成

図1に Cascade 構成を示す。Cascade 構成は3つのサイト(正サイト, 中間サイト, 副サイト)があり, 各サイトの位置関係は正サイトと中間サイト間が比較的近距离で数 km 程度, 中間サイトと副サイト間が遠距離で数百 km になっている。また, それぞれのサイト

にはストレージ(正ストレージ, 中間ストレージ, 副ストレージ)が配置される。ストレージ間にはデータ送受信の通信線が接続される。正サイトおよび副サイトには正ホスト計算機と副ホスト計算機がそれぞれ配置される。各ホスト計算機はそれぞれのサイトにあるストレージに接続される。

ストレージ間にはリモートコピーが動作する。リモートコピーとはストレージが持つデータコピー機能で、管理サーバから制御される。また、リモートコピーの種別は正ストレージ、中間ストレージ間が同期コピー、中間ストレージと副ストレージ間が非同期コピーである。ここで、同期コピー、非同期コピーは以下の特徴を持つ。

(1)同期コピー：

ホスト計算機がコピー元のストレージに対し発行したデータの書き込み完了報告を、コピー先ストレージにコピーが完了するまで、ホスト計算機に送信しない方式。特徴はコピー先でデータロスがない反面、長距離データコピーができないこと。

(2)非同期コピー：

ホスト計算機がコピー元のストレージに対し発行したデータの書き込み完了報告を、コピー処理とは非同期に、ホストに送信する方式。特徴は長距離データコピーが可能な反面、データロスが発生すること。

通常時の Cascade 構成では正ホスト計算機上

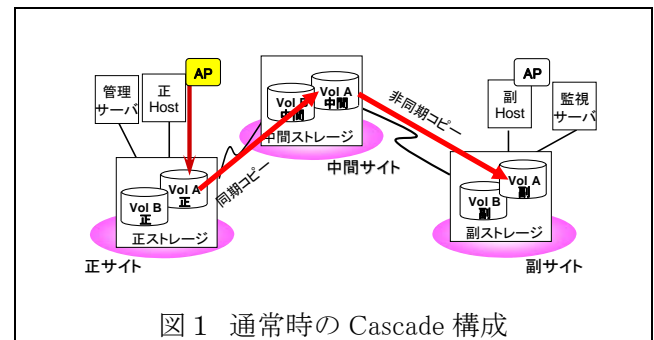


図1 通常時の Cascade 構成

でアプリケーション（図では AP）が動作し、ア

Applying High Availability Method to 3 Data Center Cascade Replication

[†]Wataru Okada, Nobuhiro Maki, Masahide Hosouchi, Yuri Hiraiwa, Masahide Sato, Systems Development Laboratory, Hitachi Ltd.

[‡]Katsuhisa Miyata, Software Division, Hitachi Ltd.

アプリケーションのデータは正ストレージに格納される。正ストレージは自装置に書き込まれたデータを中間ストレージにコピーする。また、中間ストレージは自装置に書き込まれたデータを副ストレージにコピーする。

正サイト被災時の Cascade 構成では図 2 に示すように正ホスト計算機上のアプリケーションを副ホスト計算機上に移行し、災害から最遠の場所で動作させる。アプリケーションの移行と平行して、可能であれば中間ストレージのデータを副ストレージに転送する。上記の転送が不可能であれば、副ストレージに格納されたデータが移行したアプリケーションのデータになる。また可能であれば可用性向上のため、副ストレージは中間ストレージに向け非同期コピーを通常時とは逆向きに動作させる。

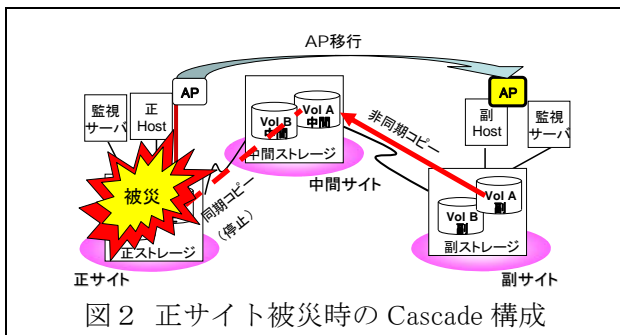


図 2 正サイト被災時の Cascade 構成

2.2 Cascade 構成の課題

図 2 に示したように正サイト被災時の Cascade 構成は通常時とコピー構成が異なる。すなわち、被災時の Cascade 構成は副ストレージからの非同期コピーのみであり、この場合コピー先でデータロスが発生しうる。上記は Cascade 構成を災害回復用途と想定するためである。正ホストの障害時や災害の規模によっては、正ストレージが動作する場合があります。その場合は本来副サイトから回復する必要はない。

3. Cascade 構成におけるデータ可用性維持方式

本方式は Cascade 構成における正ホスト計算機の障害を対象とした回復方式である。図 3 に正ホスト障害時における本方式の構成を示す。本方式は、従来構成の他に、中間ホスト計算機と、通信路を回復用に新規に設ける。上記により、中間ホスト計算機で障害前と同一コピー構成を再構築する。

具体的には、障害後アプリケーションは中間ホストで動作する。コピーの構成は中間ストレージから正ストレージへの同期コピーと、正ス

トレージから副ストレージへの非同期コピーからなる。通常構成から回復構成に移行する手順は以下のとおり。

- (1) 正ストレージから中間ストレージへの同期コピーを逆転する
- (2) 正ストレージ副ストレージ間で通信路を確立する
- (3) 中間ストレージから副ストレージへの非同期コピーを正ストレージ基点としたものに移行する
- (4) アプリケーションを正ホスト計算機から中間ホスト計算機に移行し、業務の回復手続きを行う

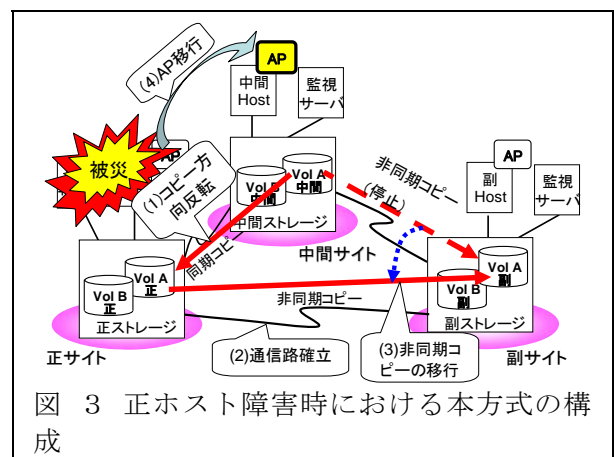


図 3 正ホスト障害時における本方式の構成

障害前後でコピー構成を比較すると、同一であることがわかる。上記の手続きにより、正ホスト計算機に障害が発生したとしても、最新データのある中間サイトから通常構成と同様のコピー構成が可能で、結果として、高い可用性の維持が可能となる。

4. さいごに

本報告では Cascade 構成のデータ可用性を維持するため、従来は利用していない中間ホスト計算機に業務を移行し、コピー構成を障害前の形態と同様になるように変更する。今後は、本方式の有効性を詳細に検証するため、実機による評価を実施する予定である。

参考文献

- [1] Rudolph, C.G.,"Business continuation planning/disaster recovery", Communications Magazine, IEEE, Volume 28, Issue 6, PP.25-28, June 1990