

Storage Fusion : ログ同期ディザスタリカバリ方式の性能評価

鈴木 芳生[†] 渡辺 聡[†] 水野 和彦[†] 藤原 真二[†] 河村 信男[‡][†] (株)日立製作所中央研究所[‡] (株)日立製作所ソフトウェア事業部

1. はじめに

911 テロ以降、災害であってもビジネス(業務)を継続させることの重要性が再認識されている。ビジネス継続ができないことによる機会損失は大きく、例えば、金融業界では時間あたり数百万ドルもの損失が生じると言われている。

そのため、米国では社会全体への影響が多い金融業に関しては、法規制を設ける動きもある。同様に、日本でも経済産業省によってビジネス継続のためのガイドラインが作成されている。さらに、国際的にも英国 BS25999[1](旧 PAS56)等を土台として ISO における標準化が予定されている。このような背景から、システムを二重化し、平常時から正サイトのデータを遠隔地に設置した副サイトにコピーしておき、災害時には副サイトでビジネスを継続するディザスタリカバリ(DR)技術が注目を集めている。

2. DRシステムの課題

日本では地震など広域災害への対応が必要となることから、都市間 DR、すなわち、サイト間の距離が数百 km 以上離れた構成での運用が必要となる。金融業などデータ欠損が大きな損失を及ぼす場合には、この都市間 DR においても、災害時のデータ欠損がないことが大きな価値となるが、同時に、オンライン性能の確保、すなわち、副サイトへのコピー処理が正サイトで行うオンライン処理に与える影響を最小限にしなければならない。

さらに、実際に DR システムを運用していくためには、日々の運用が容易である必要がある。すなわち、性能設計や性能チューニングが容易に行える必要がある。そのためには DR システムが、転送遅延や AP 負荷といった環境因子の変動に対して高い性能安定性を有する必要がある。

そのため、本論文では、データ欠損なしとオンライン性能確保の両立可能な DR システムを提案すること、及び、広域災害を想定した遠距離構成において、評価を行い、提案方式の性能安定性を確認する事を目的とする。

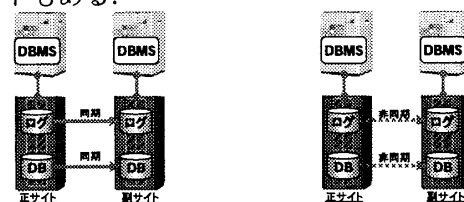
StorageFusion: Evaluation of Log-sync-transfer DR System

[†] Yoshio Suzuki, [†] Satoru Watanabe, [†] Kazuhiko Mizuno,[†] Shinji Fujiwara, [‡] Nobuo Kawamura[†] Hitachi, Ltd., Central Research Laboratory[‡] Hitachi, Ltd., Software Division

3. DBのディザスタリカバリ方式

3.1 ストレージ主体の方式

ストレージ主体の方式は、ストレージのリモートコピー(RC)機能を用いて、副サイトへの転送を行う(図 1)。すなわち、DBMS のログ・DB とも同期/非同期 RC によって転送する。同期 RC を用いる同期転送方式は、災害時にもデータ欠損がないが、平常時には同期転送の影響でオンライン性能が低下する。一方、非同期転送方式は、平常時のオンライン性能低下は少ないが、災害時にはデータ欠損なしの保障ができない。また、ストレージ主体の方式は、転送処理がストレージで行われるため、オンライン処理に多くのリソースを割当てられる、あるいは、平常時は副サイトのサーバリソースの稼働が不要というメリットもある。



(a) 同期転送方式

(b) 非同期転送方式

図-1 ストレージ主体の DR システム

3.2 DB・ストレージ連携方式

ストレージ主体の方式では、ログ・DB 共が転送対象となるため、転送量が多くなる。都市間 DR においては、平常時から数百 Km の回線の維持コストの負担が必要となるため、転送量の削減は重要な課題である。

我々は、同期 RC でログのみを転送し、副 DB はログ適用によって回復するログのみ同期転送方式を開発した(図 2)[2]。ログ適用処理は正サイトのオンライン処理に比べて負荷が軽いことから、副サイトのログ適用処理はアプライアンス(DB ヘッド)で実施する。DB ヘッドの利用によって、平常時は高価なサーバの稼働を不要とすることが可能となる。

本論文の目的は、広域災害を想定した構成で、災害時の欠損なしを実現する DR システムの性能評価を行う事である。そのため、同期 RC を利用した、同期転送方式とログのみ同期転送方式を対象として性能評価を行った。

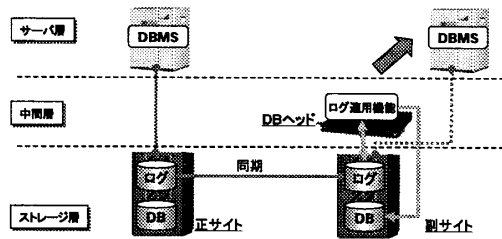


図-2 DB・ストレージ連携 DR システム

4. 評価

4.1 評価方法と評価環境

性能評価は図 3 に示す実験環境で行った。すなわち、DB サーバとストレージのセットを正/副サイトとみなし、正副ストレージの間を回線シミュレータで接続した。正 DB サーバには AP サーバを接続し、この AP サーバから、TPC-C をベースにした販売管理モデルを用いて負荷をかけることで評価実験を行った。

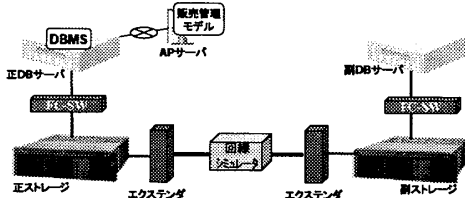
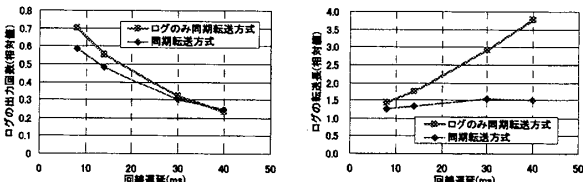


図-3 評価環境

4.2 評価 1: オンライン性能の評価

回線シミュレータの転送帯域幅を固定し、転送遅延をパラメータとして実験を行い、転送遅延がオンライン性能に及ぼす影響を評価した。図 4 に、ログの出力回数とログの転送長を示す。



(a) ログ出力回数

(b) ログ転送長

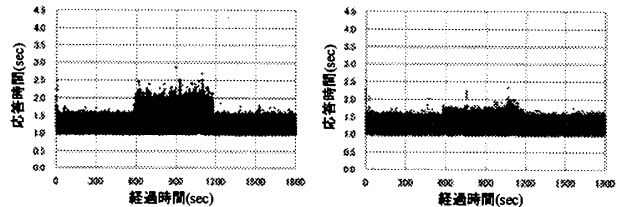
図-4 ログの出力回数と転送長

図 4(a)より、同期転送方式/ログのみ同期転送方式とも、転送遅延が大きくなると、ログ出力回数は減少している。一方、転送長については、同期転送方式では転送遅延に関わらず、一定の転送長であるのに対して、ログのみ同期転送方式では、回線遅延が大きくなると、転送長も大きくなっている(図 4(b))。ここで、ログ出力回数と転送長の積であるログの転送量、すなわち、正 DBMS のオンライン性能となる。そのため、ログのみ同期転送方式は、転送遅延が大きくなり、ログの出力頻度が制限されるような場合であっても、転送長を大きくすることでスループット(オンライン性能)を確保することができる。

4.3 評価 2: 性能安定性の評価

入力負荷が動的に変動する環境での評価を行う事で、DR システムの性能安定性を評価した。入力負荷は急激なピークが生じるようなスケジュールで投入した。すなわち、定常的にベース負荷を流しておき、中間期間に追加負荷を加える事で、負荷集中期間を作りだす事にした。

図 5 に負荷集中期間の合計負荷として、ベースの 6 倍の負荷を与えた場合の、応答時間を示す。図より同期転送方式では、負荷集中期間に応答時間が大きくばらついているのに対し、ログのみ同期転送方式では、ばらつきが小さくなっていることがわかる。そのため、ログのみ同期転送方式は、AP 入力負荷変動といった環境の変動に対してもロバストであり、性能安定性が高いことが確かめられた。



(a) 同期転送方式

(b) ログのみ同期転送方式

図-5 性能安定性

5. おわりに

広域災害を想定した遠隔構成において、災害時の欠損なしと平常時のオンライン性能確保を両立する方式として、ログのみを同期リモートコピーで転送し、副 DB は DB ヘッドでのログ適用で回復するログのみ同期転送方式を提案し、回線シミュレータ環境で性能評価を行った。

評価の結果、ログのみ同期転送方式は、転送遅延の影響でログの出力頻度が制限される場合であっても、転送遅延に応じてログの転送長を大きくすることで、オンライン性能を確保可能なことが確かめられた。また、入力負荷が動的に変動する環境で評価を行った結果、ログのみ同期転送方式を利用した場合は、応答時間のばらつきが小さく、高い性能安定性が得られる事が確かめられた。

謝辞 本研究の成果の一部は、文部科学省リーディングプロジェクト e-Society 基盤ソフトウェアの総合開発「先進的なストレージ技術」の支援により得られたものである。

6. 参考文献

- [1] British Standards Institution. BS25999: Business Continuity Management, 2006.
- [2] 日立製作所. ログのみ同期転送で通信コストを削減する高信頼 ディザスタリカバリ技術. はいたつく, Vol. 8, pp. 15-16, 2005.