

2H-08 プロセスマイグレーション技術の リレーショナルデータベースへの応用

村田 明文¹、白木原 敏雄²、平山 秀昭³、小石 誠¹、蔵野 政行¹

¹(株)東芝 デジタルメディア機器社 ²(株)東芝 研究開発センター

³(株)東芝 コンピュータ&ネットワーク開発センター

1. はじめに

コンピュータが社会の基盤として重要になるにつれて、サーバシステムの高可用性への要求がますます高まってきている。

従来、ハードウェア故障を想定したサーバシステムの高可用性のために、「フォルトトレラント (FT) 機を用いる構成」と「クラスタ構成」の2種類の構成が一般的に用いられてきた。しかし、FT 機は特殊なハードウェアを必要としシステムが非常に高価になる為、クラスタ構成を用いる方法がより広く使用されている。一方、クラスタ構成でリレーショナルデータベース (RDB) を使用する場合は、障害復旧時に RDB のロールバック及びロールフォワード時間が長くなるという問題がある。

このクラスタ構成で RDB を使用する場合の問題点 (従来技術の問題点) をプロセスマイグレーション技術 (実行途中のプロセスを再現する技術) を応用することで解決したので、その手法について報告する。

2. 従来技術とプロセスマイグレーション技術

従来技術とプロセスマイグレーション技術の障害復旧処理における違いについて記述する。

従来技術を用いる方法では、稼働計算機に障害が発生し待機計算機に実行中のプロセスが引き継がれる場合に当該プロセスを最初から再起動する。当該プロセスを障害発生時点の状態に復帰させる為に必要なデータは、共有ディスク上に記録され、障害復旧の為に再起動時に利用される。

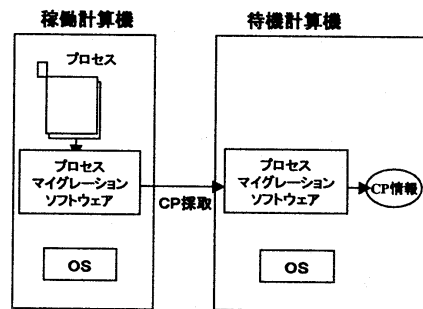


図1 プロセスマイグレーション技術

次に、プロセスマイグレーション技術を用いる方法を記述する。プロセスマイグレーションソフトウェアは図1に示すように、稼働及び待機の両方の計算機上で動作し、定期的にプロセスマイグレーションに必要な情報 (チェックポイント (CP) 情報) を採取する。稼働計算機に障害が発生した場合、待機計算機のプロセスマイグレーションソフトウェアは CP 情報をもとに当該プロセス状態を CP 採取時へ戻し再開始 (CP ロールバック) を行う。

3. プロセスマイグレーション技術の RDB への応用

RDB を使用するシステムを従来技術を用いてクラスタ構成にする場合、障害復旧時間が長い (RDB の再起動処理からロールフォワード及びロールバックに数分~数十分必要) とい

"Process migration mechanism for Relational Data Base System."

Akifumi Murata, Makoto Koishi, Masayuki Kurano
TOSHIBA Corporation
Digital Media Equipment & Services Company
Toshio Shirakihara
TOSHIBA Corporation
Corporate Research & Development Center
Hideaki Hirayama
TOSHIBA Corporation
Computer & Network Development Center

う問題がある。ロールフォワード及びロールバックの時間は、RDB プロセスの再起動を行う限り避けられない問題であり、従来技術では回避することが出来なかった。

筆者らはこの問題を解決するため、プロセスマイグレーション技術を RDB に応用するソフトウェアを開発した。このソフトウェアによって RDB プロセスはロールバック及びロールフォワードの必要が無くなり、障害復旧時間を数秒～数十秒とすることが可能になった。

4. プロセスマイグレーションの RDB での問題点と解決手法

次に、プロセスマイグレーション技術を RDB プロセスに応用する際に発生した問題点とその対策について記述する。

前述のように、プロセスマイグレーション技術では、障害復旧時にプロセスの CP ロールバックを行う。RDB プロセスが CP ロールバックされる場合の問題点について図 2 を用いて説明する。

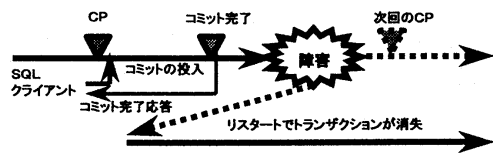


図2 プロセスマイグレーションのRDBでの問題点

CP 採取から次の CP 採取までの間に SQL クライアントプロセスがコミットを発行し確認応答を受けた状況を想定する。この状況で、次の CP 採取を行う前に、計算機障害が発生すると、コミット投入前の CP まで RDB プロセスがロールバックされてしまう。つまり、これはコミットされたトランザクションの消失を意味している。RDB の最も重要な機能の一つは、トランザクション性の保証であるが、このままではトランザクション性を保証できない。

筆者らのこの問題に対する解決手法につい

て図 3 を用いて説明する。

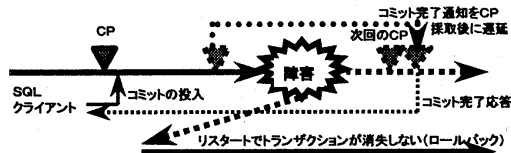


図3 プロセスマイグレーションのRDBでの問題点の解決

RDB プロセスから SQL クライアントプロセスへのコミット応答を次の CP 採取まで遅らせるようにした。これにより、SQL クライアントプロセスからみてコミット応答があり、その直後に RDB プロセスに障害が発生した場合でも、確実にそのトランザクションが保証される事を可能とした。

また、コミット応答を遅延させる処理はレスポンスを最大で CP 採取間隔時間遅らせるが、トータルスループットの減少は数～十数%と実用上大きな影響は無いことを確認した。

5. まとめ

従来技術で RDB システムをクラスタ構成とした場合の問題点を解決するため、RDB プロセスをマイグレーションするソフトウェアの開発を行った。これにより、従来技術で大きな問題となっていた障害復旧時間を飛躍的に短縮することが可能となった。

プロセスマイグレーション技術は、動的負荷分散システムへの適用など、様々な可能性を秘めており、さらなる応用を検討し、実装して行く予定である。

6. 参考文献

- [1]村田・蔵野・末永 第56回情報処理学会全国大会「分散ノード連携技術による高信頼システムの実現」
- [2]白木原・平山・佐藤・金井 第54回情報処理学会全国大会「高信頼ミドルウェア ARTEMIS の概要とチェックポイント生成方式」