

高速トランザクション処理のための  
リカバリ情報取得方式

4 R - 2

北川 政人<sup>1)</sup> 萩原 敏幸<sup>1)</sup> 藤田 裕夫<sup>2)</sup>  
1) NTT情報通信網研究所 2) NTTソフトウェア(株)

1. はじめに

近年のオンライントランザクションシステム(以下OLTPと略記)では、トラフィックの増大により処理の高速化及び高信頼化が最重要課題とされている。これらの課題を解決させるため、処理の高速化ではCPU性能の向上等のハードウェアに関する性能向上とプログラムのダイナミック・ステップ(以下DSと略記)の削減や磁気ディスク等の補助記憶装置へのI/O回数削減等のソフトウェアに関する性能向上の検討が進められている。また、高信頼化では、システムの自動系切り替え機能、障害機器を切り離す縮小運転機能、ファイル回復機能等のシステムリカバリ機能が提供されている。

本稿では、ファイル回復機能(以下リカバリ機能と記述)において、ファイル回復のために必要な情報を取得するためのI/O回数を削減することにより、高速かつ高信頼のトランザクション処理を実現する方法について報告する。

2. 従来方式における問題点

OLTPではトランザクションの処理結果を保証するため、トランザクション毎に更新ログ等をリカバリ情報としてジャーナルファイルに取得し、障害発生時はそれをもとにシステム内のファイルの状態を矛盾無く回復させる機能が提供されている。

これらの機能を実現するため、トランザクションによるファイル更新(FILE\_WRITE)はまずメモリ上の更新データバッファにおいて当該更新情報を書き込み後、トランザクション終了時(TR\_END)に以下の手順で処理される必要があった。(図2.1参照)

手順①: トランザクション内の更新処理でメモリ上に展開された更新情報をリカバリ情報としてDK上のジャーナルファイルに書き込む

手順②: 更新情報をDK上のファイルに書き込む

手順③: 該トランザクション処理の完了をメモリ上のトランザクション管理テーブルへの書き込み後、リカバリ情報としてDK上のジャーナルファイルに書き込む

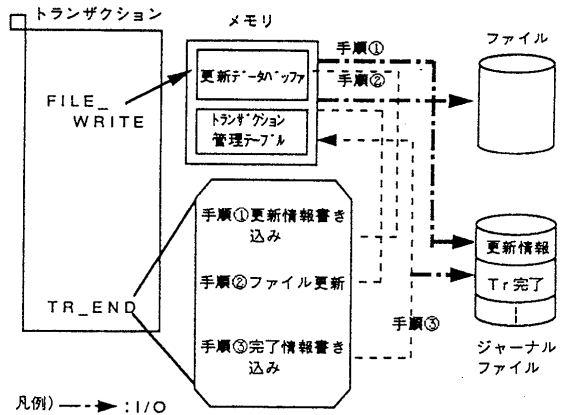


図2.1 従来のジャーナル取得処理の流れ

上記のリカバリ情報取得方式では、ファイルの内容を更新するためのI/Oに加え、更新情報をジャーナルファイルへ書き込むためのI/Oと、該トランザクションの完了を書き込むためのI/Oを前後で行なわなければならないため、ひとつのトランザクションにつきI/Oが3回となり、通常処理のスループットがI/Oネックにより低下するという問題があった。

3. トランザクション処理の高速化

本方式では、更新情報をジャーナルファイルへ書き込む際、該トランザクション終了直前に完了した全トランザクションの完了情報を同時にDKへ書き込むことによりトランザクション当りに発行するI/O回数を削減し、トランザクション処理の効率化を図った。

(図3.1参照)

Recovery Information Control Method  
for High Throughput on OLTP Systems

Masato KITAGAWA<sup>1)</sup>, Toshiyuki HAGIWARA<sup>1)</sup>, Yasuo FUJITA<sup>2)</sup>

1) NTT Network Information Systems Laboratories, 2) NTT Software Corporation

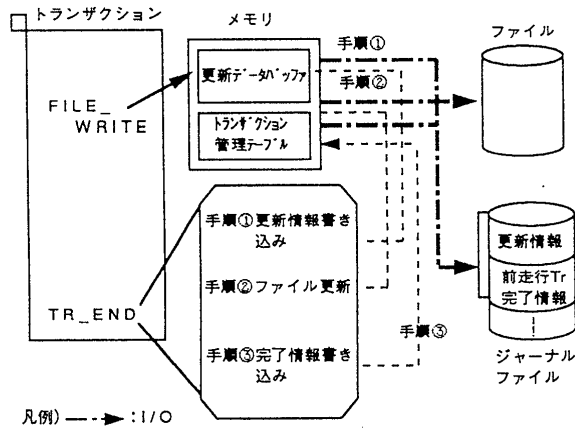


図3.1 本方式でのジャーナル取得の流れ

4. 本方式の特長と問題点

本方式では自トランザクションの完了は次に走行するトランザクションの手順①でジャーナルファイルに取得されるため、次のトランザクションが走行しない限り自トランザクションの完了がリカバリ情報に反映されない。

つまり、トランザクション処理は完了してもリカバリ情報では次トランザクションが走行するまでは未完状態であり、その期間はトランザクション完了状態が常に遅延して管理されることになる。(図3.2参照)

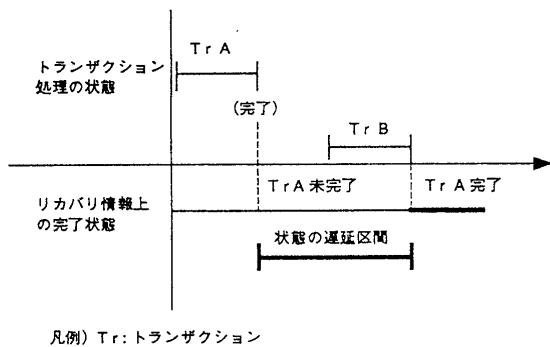


図3.2 リカバリ情報での状態遅延

このことから、従来方式ではシステムリカバリ時に完了しているトランザクションに対するリカバリ処理は行っていなかったが、本方式ではトランザクション完了状態の遅延区間でシステム異常が発生した場合、システムリカバリ時に処理が完了しているトランザクションでもリカバリ情報上

は未完状態となっているためシステムリカバリで再度処理を行うというオーバーヘッドが生じる。(図3.3参照)

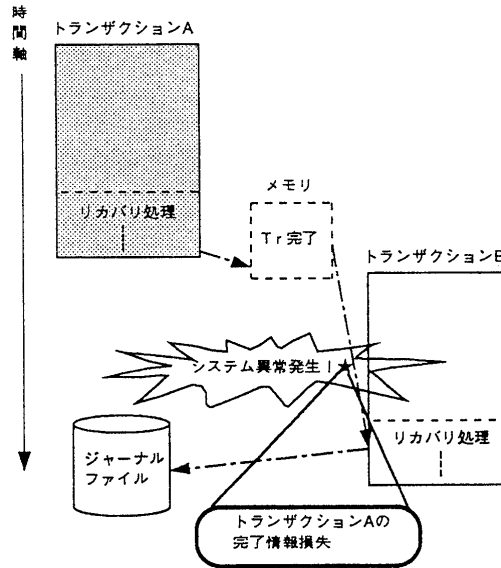


図3.3 リカバリ時オーバーヘッドとなるトランザクション

しかし、このオーバーヘッドにかかる処理時間は、全体のシステムリカバリ時間と比して非常に小さいため、トータルとしての処理高速化に対する効果の方が有益である。また、トランザクションの整合は十分とれるため、システムの信頼性を損なうことなく、トランザクション処理の高速化を実現できる。

5. おわりに

ここまでOLTPにおいて、通常処理でのリカバリ情報取得のI/Oを削減することにより、トランザクション処理の性能を向上させる方法について述べてきた。

今後は、処理の高速化の観点から複数トランザクションのジャーナルに対するバッファリング効果の向上及びトランザクション管理テーブル等のシステム共通資源に対する排他制御によるスループット低下の防止等を主眼に更に検討を進めていく予定である。