

# トランザクションAPのオンライン保守方式の評価

5D-10

白石 正裕、浜口 三明、富田 清次、西原 琢夫

NTT情報通信研究所

## 1 はじめに

OLTPシステムにおいて、データベース（以下、DB）再構成を行う場合、DB再構成の種類によっては、アプリケーションプログラム（以下、AP）の入れ替えを実施する必要がある。従来このような場合にはトランザクション処理を中断する必要がある、いわゆる24時間運転の妨げとなっていた。

本稿では、再構成後のDBの切り替えと同期してトランザクション処理APの切り替えをトランザクション処理を中断せずに実施する方式を提案する。また、トランザクション処理が再構成処理時間に与える影響について、シミュレーション評価を行うことによって本方式の有効性を明確にした。

## 2 方式の概要

### 2.1 同期した切替の必要性

テーブル削除やカラム属性変更等のDB再構成を行う場合、該DBにアクセスするトランザクション処理APについても変数の型や構造体の修正を必要とする。オンライン中に無中断で再構成を実施するには、DBの切り替えと同期してDBをアクセスするAPを切り替える必要がある。

### 2.2 方式の特徴

DBの切り替えと同期したトランザクション処理APの切り替えをトランザクション処理を中断せずに実現する方式を図1に示す。以下に本方式の特徴をまとめる。

#### (1) 複製DBによる再構成の実施

DBの複製を行い、複製DBに対して再構成処理を実施する。これにより、トランザクション処理に影響を与えることなく、再構成処理を行うことが可能である。

#### (2) データ更新履歴情報の利用

OLTPシステムでは障害時の復旧を目的として、データ更新履歴情報（以下ジャーナル）を取得している。このジャーナルを利用して複製DBを最新状態に更新する。

ジャーナルは再構成前のDBの更新情報であるため、複製DBの反映においてはデータの型変換を行う必要があり、該変換は入れ替えAPで実施する。

#### (3) ジャーナル反映用キューの設置

ジャーナル反映用キュー（FIFOキュー）を設ける。キューイング中のジャーナル数によって、新旧DBのデータ更新内容の一致性の確認が可能であり、ジャーナル数が0になったらDBとAPを切り替える。

#### (4) 新APの事前ローディング

切り替えに先だって新APを事前にローディングする。これにより、AP切り替え時のサービス中断時間が短縮できる。

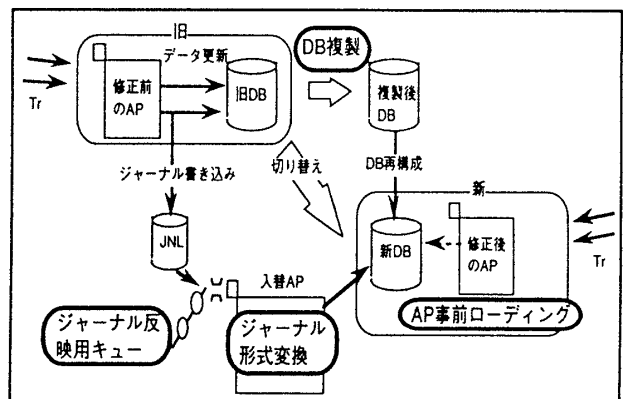


図1：DB再構成を伴うAP切替の処理概要

### 2.3 システム構成への適用

本方式を適用する場合のシステム構成として、シンプル構成とデュプレックス構成が考えられる。シンプル構成の場合は、CPU、メモリ、ディスクそれぞれ単一構成であり、1システム内で本方式を適用することができ、実現上簡易である。しかし、1システム内でトランザクション処理と再構成処理を並行して行うことにより、CPUに負荷を与えることになる。

一方デュプレックス構成の場合は、現用系側でトランザクション処理、待機系側で再構成処理を行うことが可能であるため、CPUへの影響はない。ただし、現用系、待機系の系間で新旧DBの切り替えとAPの入れ替えの同期をとるため、プロセッサ間通信を用いたデータの受渡しを行う必要があり、処理が複雑になる。

### 2.4 処理シーケンス

図2にシンプルックス構成における本方式の処理シーケンスを示す。

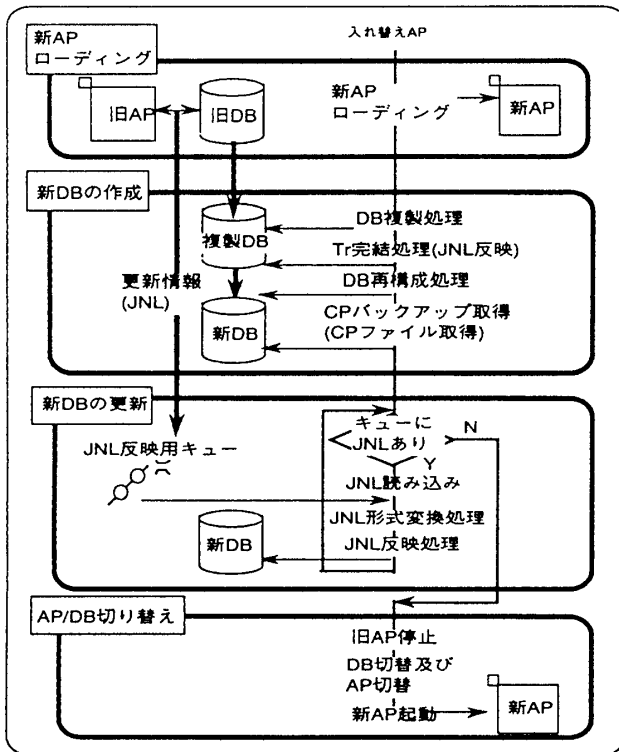


図2：DB再構成を伴うAP切替の処理シーケンス

## 3 シンプルックス構成における本方式の評価

### 3.1 評価の目的

シンプルックス構成において、2で述べた本方式を適用する場合、同一CPU上でトランザクション処理と並行して再構成処理を行うことにより、サービス中のトランザクション処理に負荷の面で影響を与える。

しかしながら、OLTPシステムにおいて、トランザクション処理のターンアラウンド時間に影響を与えてはならないため、OSの機能を用いて処理の実行優先度の設定を行い再構成処理よりもトランザクション処理に対して優先的にCPUを割り当てる。その場合、逆に再構成処理に対して時間を要することとなる。

従って、今回はトランザクション処理が全再構成処理時間に対してどのくらい影響を与えるかについて評価を行った。

### 3.2 評価内容

今回の方式に対して、以下の項目の評価を行った。

- ①トランザクション発生数による再構成全体時間への影響
- ②DB規模に対する再構成全体時間への影響
- ③再構成処理にかかる実行命令数(以下、DS値)による再構成全体時間への影響

### 3.3 評価結果

上記について、評価を行うために、我々が設計の対象としているトランザクション処理ノードをモデルとして、シミュレーション評価を行った。

#### (1) 測定パラメータ

本方式の評価における各種設定を表1に示す。

表1：各種設定

項目	測定値
CPU処理能力	8 (MIPS)
DK I/O時間	3 (msec)
書き込みブロック長	256 (byte)
1 TrあたりのDS値	35 (kstep)

#### (2) 評価結果

シミュレーション評価プログラムにて評価した結果について、トランザクション発生数に対する全再構成処理時間の関係(再構成対象DBレコード数別)を図3に示す。トランザクション発生数に対する全再構成処理時間の関係(再構成処理DS値別)を図4に示す。

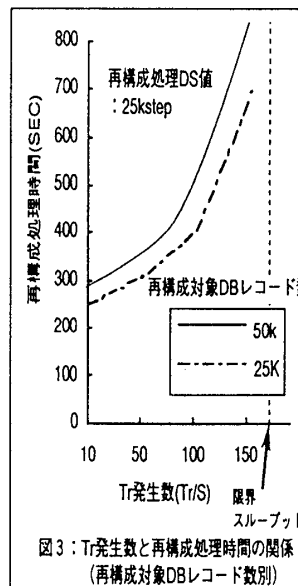


図3：Tr発生数と再構成処理時間の関係(再構成対象DBレコード数別)

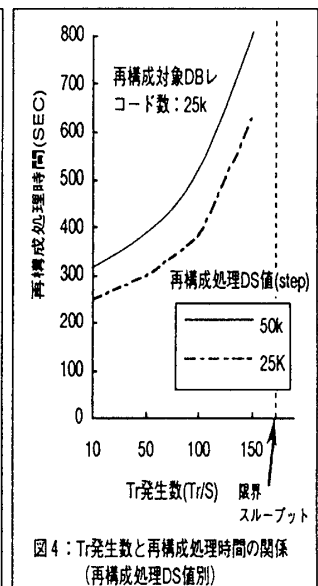


図4：Tr発生数と再構成処理時間の関係(再構成処理DS値別)

評価した結果、以下のことがわかった。

- ・再構成処理時間は、一般的な設計条件範囲では、CPU処理能力に対する限界スループットの約50%の負荷までなら上に上昇し、それ以降は急激な上昇となる。よって、本方式による再構成処理を行う場合、少なくともCPU処理能力の50%以上必要になることがわかった。

### 4 おわりに

DB再構成に同期してトランザクション処理APの入れ替えをサービス無中断で実施する方式について、シミュレーション評価を行うことにより、システム適用領域を明確化した。今後更に実際のシステム上で本方式を適用し、運用可能であるか評価する予定である。