

## 手順内包型DBMSによるオン中保守方式\*

2F-7

堀口 恭太郎 梅本 佳宏 中村 仁之輔

NTT情報通信網研究所

## 1. はじめに

24時間無中断で稼働するネットワークの構成要素としてのDBMSではDB再構成をサービスを継続しながら行うオン中DB再構成が重要となる。

現状、DBオン中保守としてはおもにデータのロード/アンロード、また一部のDBMSでは再編成が実現されつつあるが、DBの定義(スキーマ)を変更する再構成や、データ実体のファイル内での移動を伴う再編成の処理は、一般的にサービス停止状態で行う必要がある。

ネットワークに限らずOLTPシステムではオン中DB再構成の機能を充実させることはサービスの無中断性、運用の柔軟性の向上という点で有用であると考えられる。

そこで本報告では、リレーショナルDBMSにおいて、カラム追加、削除、属性変更、一括更新などのDB再構成をシステム稼働状態で行う、オン中DB再構成の機能を実現する機構とその機構上での再構成の方式について提案する。

## 2. 無中断DB再構成

無中断DB再構成を、ここでは「スキーマ切り替え(世代遷移)をAPから見ておおむねトランザクションの平均到着間隔より短い時間以内で行う」とこととする。オンライントランザクション処理を行うシステムでは通常の方法ではこの時間内にDB再構成を終了させることはできない。

そこでここではスキーマの切り替えをレコード単位で行うことでオン中DB再構成の実現を行う方式について提案する。

## 3. アクセス機構-手順内包アクセス機構

## 3.1 検討課題

オン中DB再構成を実現するにあたって以下の点が課題として挙げられた。

- (1) サービス無中断で再構成を行うと、旧世代APのトランザクションのすべてが終了する前に新世代APのトランザクションがスタートし、新/旧世代APの同時走行の可能性がある。しかし走行中の旧世代APは再構成前のDBを、また走行し始める新世代APは再構成後のDBを同時に要求するため、DBMSはこの二世代のDBの同時アクセスを処理する必要がある。
- (2) メモリデータベースなどでテーブル領域に余裕がないような場合ではワークテーブルを作成することができない。そのような場合でも再構成ができるようにするためワークテーブルを使わないようにするならば、両世代のAPが同じテーブル実体をアクセスできるようにする必要がある。

つまり、一つのデータをAP世代に応じたイメージで見せることが必要になる。

## 3.2 手順内包アクセス機構

以上の問題を解決するために図-1のようなアクセス機構を提案する。

DBMSはアクセスメソッド(手順)を内包し、それぞれの世代のAPはこのメソッドを通してその世代に応じた表現でデータを操作する。

またデータアクセス機構を実データアクセス部と抽象データアクセス部とに分離し、テーブルの記憶構造の変更とAPに見えるデータ表現の変更を独立に行えるようにすることでワークテーブルを用いないDB再構成を実現する。

## 4. 再構成シーケンス-2段階再構成

本方式では再構成作業は基本的にワークテーブルを用いず、進行はレコード単位とする。そのため

\* On-Line Maintenance of Database Schema on Encapsulated DBMS.

Kyotaro Horiguchi, Yoshihiro Umemoto,  
Jinnosuke Nakamura  
1-2356 Take Yokosuka-Shi Kanagawa 238-03, Japan

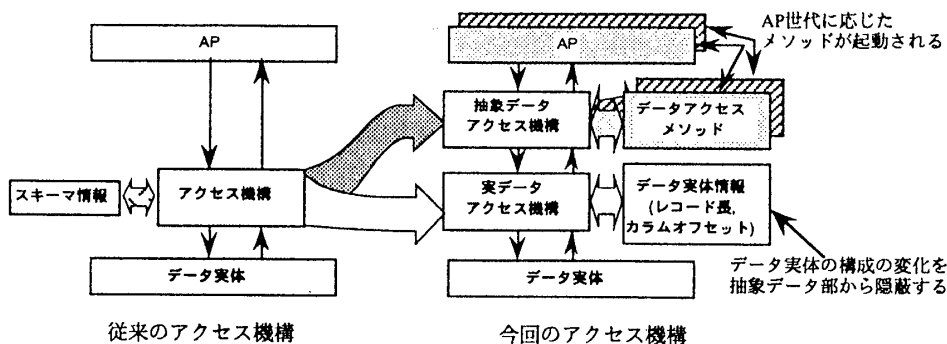


図-1 基本方式

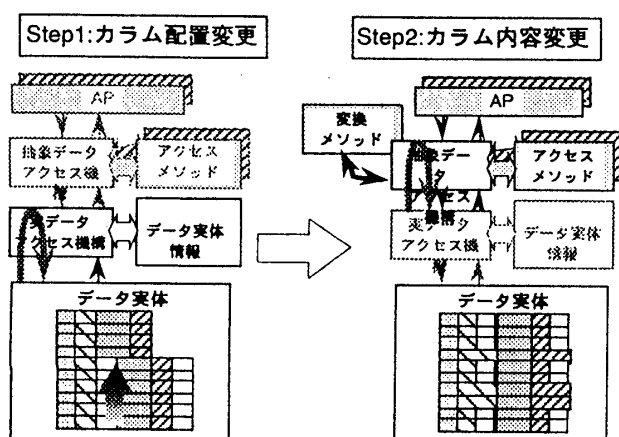


図-2 2段階再構成

に図-1に示すような機構を用いて再構成を2つの段階に分解して行う(2段階再構成:図-2).

Step1:(図-2左)

この段階では再構成の際のデータ変換がレコード領域内で行えるようにするために、レコードの物理的構造の変更(カラム配置変更)を行う。

配置変更は領域拡大となるので、テーブル後方から順次作業を行っていく。

この作業は2段アクセス機構の実データアクセス機構が行うが、この変化は抽象アクセス機構から上の部分には影響を及ぼさない。

Step2:(図-2:左)

Step1終了後に今度は抽象データアクセス機構がレコードアクセスの際にデータ型/内容の変更(カラム内容変更=世代遷移)を変換メソッドを用いてオンデマンドで行っていく。

この作業自体も2段階で進行する。

Step2前半

これは、まだ旧世代APのトランザクション

が残っている状態である。

レコードが世代遷移してしまうと旧世代APによる更新が不可能になるので世代遷移は必要最低限、新世代APによる更新が行われた時のみに行う。旧世代APによる世代遷移後のレコードの参照は世代遷移時のBIログを用い、更新を行ったときはトランザクションをアポートさせる。

このためにトランザクションIDを用いたマルチバージョン管理<sup>(1)</sup>を行う。

Step2後半

旧世代APの全トランザクション終了後は、DBMSが全てのレコードに対して世代遷移を順次行いDB再構成が完了する。

さらにこの後で必要なら余った領域を詰める処理をStep1と同様の手順で行う。

5.まとめ

2段アクセス機構とマルチバージョン管理の組み合わせによって、レコード単位での再構成世代管理を行う手法を示した。これにより、ワーク領域の確保が不要な漸進的なオン中再構成シーケンスを考案し、サービス無中断でのDB再構成の実現方法を示した。

再構成にワークテーブルを使わないこの方式は大量のワーク領域を用意する余裕のない場合が多いメモリデータベースのような用途に適している。

参考文献

(1) Jeffrey D. Ullman(大保信夫, 國井利泰訳) : データベースシステムの原理, 日本コンピュータ協会, 1985  
 (2) Steven Bobrowski 他 : ORACLE7 Server概要, Oracle Corporation, 1993