

メモリデータベースシステムにおける データ転送方式*

3 R-7

岡田 敏† 小濱 千恵‡ 板倉 一郎§

NTT 情報通信研究所¶

1 はじめに

フリーダイヤル等の通信サービスでは、その制御データの管理をデータベース(DB)システムが行っている。このようなDBシステムでは、フリーダイヤル番号と実電話番号とのマッピングのような単純な処理を、非常に高速・高スループットで実行しなくてはならない。高速化を実現するため、我々はDBをメモリに常駐化したメモリDBシステム[1][2]を研究開発してきた。

近年、通信サービスの多様化が進み、DBシステムでの処理内容が複雑化している。例えば、ディレクトリサービスでは、取得すべき通信情報を絞り込むために、JOIN, INTERSECTION, ORDER BY等の処理が要求され、また、従来の通信サービス同様、高速な処理が望まれている。

このような複数レコードに対する横断的な処理(以下、集合操作と呼ぶ)を高速に実行するため、我々はSQLよりも下位のレベルのインタフェースをDBMSに設けることを提案し、そこで使用するデータ転送方式について検討を行った。

2 システム構成

SQLではJOIN等の集合操作の処理内容を指定するだけで、具体的な処理シーケンスを指定することはできない。DBMS内の最適化機構は様々なAPに平均的に高速化されるように処理シーケンスを決定するため、特定のAPを考えたときに、必ずしも最適であるとは限らない。

特定のAPに最適化し、処理性能を向上させるためには、AP側で具体的な集合操作の処理シーケンスを記述できるインタフェースが必要である。

そこで、図1(b)に示すシステム構成を採用する。この構成では、DBMSはDBへの高速なアクセス機能のみをサポートする。そして、APが求める集合操作はAPに最適化し、AP側で実現することによって、システム全体の高速化を行う。

このシステム構成を採用する場合、DBMSからAP

へ集合操作のためのデータが大量に転送されることになる。そのため、DBMS-AP間のデータ転送の高速化を行う必要がある。また、システム全体の高速性を保つため、システム全体を極力メモリ上で動作させなくてはならない。このとき、メモリという限られた資源を有効に使用する転送方式が必要となる。

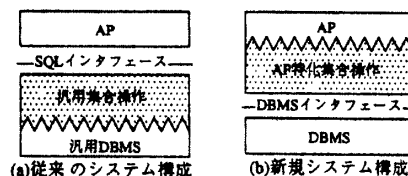


図1: システム構成

3 データ転送方式

図2(a)の従来のSQLインタフェースにおけるデータの転送(逐次転送方式)から、以下の3点に注目する。

転送単位: データ一件毎の転送、あるいは複数件のデータをまとめた転送。

転送手段: DBMSプロセスを通してのデータ転送(間接転送)、あるいはDBから直接データをコピー(直接転送)。

転送対象: データ実体の転送、あるいはデータのアドレスの転送。

以上の観点から3種の転送方式について検討する。

1. データ一括転送方式

図2(b)に示す方法で、複数件のデータをまとめて転送することによって、APとDBMS間の転送回数を削減し、転送速度の向上を図る。

2. データコピー方式

従来、APはDBMSプロセスを介して、DBへアクセスしてきたため、APとDBMSの間にはプロセス間通信が必要であった。図2(c)に示す方式は共有メモリ上に構築されたDBをAPから直接参照し、転送手段をプロセス間通信から、メモリコピーにすることによって、高速化を図る方式である。

3. アドレス参照方式

図2(d)に示す方式は図2(c)の構成でデータのアド

*Methods of Transferring Data for Memory Database Systems

†Satoshi Okada (satoshi@dq.isl.ntt.co.jp)

‡Chie Kohama (kohama@isl.ntt.co.jp)

§Ichiro Itakura (itakura@dq.isl.ntt.co.jp)

¶NTT Information and Communication Systems Laboratories

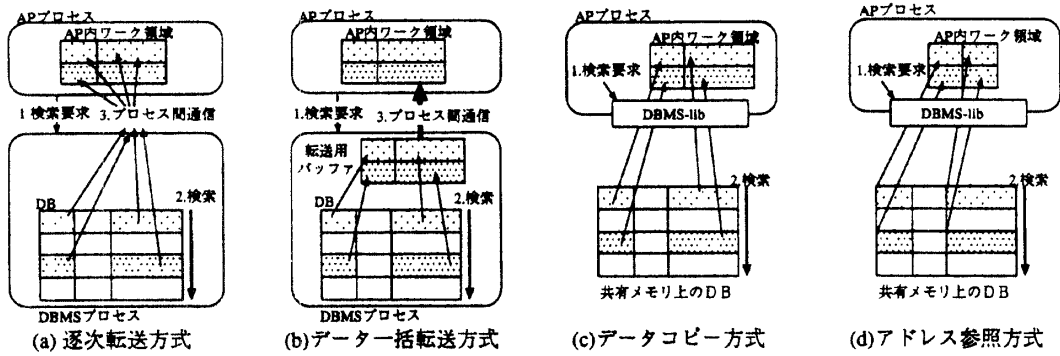


図 2: データ転送方式

ドレスを転送することによって、メモリ使用量の削減を図る方式である。AP側はアドレスしか保持していないため、データ参照時にはDB内のデータ格納情報が必要である。

4 評価

評価用プログラムを作成し、各転送方式の定量的評価を行なった。10000件のテーブルから1000件を選択し、INT,CHAR(256)の2カラムのデータをAP側に返却する場合の転送時間、参照時間、メモリ使用量を計測した。評価は512[Mbyte]のメモリを実装したSun Ultra 1上で行った。

4.1 評価結果

評価結果を表1に示す。この評価結果は逐次転送方式を100とした相対値で表している。

表 1: 評価結果

転送方式	転送時間	参照時間	メモリ使用量
逐次転送方式	100	100	100
データ一括転送方式	61.1	110	200
データコピー方式	1.95	96.2	100
アドレス参照方式	1.57	157	3.13

転送時間: 直接転送は間接転送よりも50倍以上高速である。また、データ一括転送は逐次転送に比べ4割、アドレス参照は、データコピーと比べ1割の高速化が確認できる。

参照時間: アドレス参照はそれ以外に比べ、5割増しである。これは主にDB内のデータの型情報の参照、及び型変換の時間である。

メモリ使用量: アドレス参照はINT+CHAR(256)が2アドレスサイズでよいため、1/30以下で済む。データ一括転送は転送用バッファが要するため、逐次転送の2倍必要である。

転送時間 + (参照時間 × 参照回数): 参照回数が3回以下であればアドレス参照、3回を越えればデータ一括が高速である。

4.2 考察

各方式とも利点、欠点があるため、APに応じた使い分けが必要である。各方式の特徴を以下に示す。

間接転送: APとDBMSの独立性が必要である場合。

逐次転送方式: メモリ量を考慮すると、AP側でcountなどの逐次処理を行なう場合。

データ一括転送方式: 転送量が多い場合に性能差が大であるが、メモリ量は2倍必要である。

直接転送: 性能を重視し、APとDBMSの独立性が必要でない場合。

データコピー方式: 性能では参照回数が3回を越える場合。メモリ使用量では転送データがアドレスサイズ以下。

アドレス参照方式: データコピー方式の適用範囲以外。

5 まとめ

多様化した通信サービスの高速化を目指して、AP側でAPに最適化した集合操作を行なうシステム構成を採用し、その時使用するデータの転送方式について検討、評価した。この結果、通信サービスでは直接転送が有効であり、APで行なう集合操作の性質によりデータコピー方式、アドレス参照方式を使い分ける必要があることが分かった。今後は、転送方式以外のインタフェースについても検討を進めて行く予定である。

参考文献

- [1] T.Honishi, N.Kobayashi and J.Nakamura: Design and Implementation of an Enhanced Relational Database Management System for Telecommunication and Network Applications, in *Proceedings of Pacific Telecommunications Council Eighteenth Annual Conference*, pp.698-703, 1996.
- [2] 芳西 崇, 小林 伸幸, 中村仁之輔, 田中 豪: リアルタイム DBMS 構築技術, *NTT R&D*, Vol.45, No.1, pp.27-32, 1996.