

データベース管理システムにおける光ディスクライブラリ装置のサポート方式

7Q-5

伊藤 真一、長濱 芳寛、杉山 守  
(NTT 情報通信処理研究所)

1 はじめに

近年、マルチメディア情報処理分野の発展に伴い、文字・数値データ以外に、図形・画像等の長大データのデータベース化が望まれている。長大データを文字・数値データと共にリレーショナルデータベース管理システム(DBMS)で一元管理する方式は先に提案した<sup>1,2)</sup>。しかし、長大データのデータベース化ではデータ量が多いため(数十K~数MB)、ファイルコストが高くなるという問題がある。長大データのアクセス頻度が低い分野では、長大データの記憶媒体としてライブラリ型の記憶装置を利用すればファイルコストを抑え、大容量のデータベースを経済的に構築することが可能となる。

本稿では、大容量、低コストの新しい記憶装置である光ディスクライブラリ装置[OLU]をデータベースの格納装置として利用する場合の問題点と解決方式及び、その評価結果について報告する。

2 OLUの概要

OLUは、光ディスク[OD]を格納するセル、データのリード・ライトを行う光ディスクドライブ[ODD]、セル~ODD間でODを運搬するアクセッサ及び、制御装置[ODC]等で構成される(図1)。通常、ODDは2~4台で、ODは数十枚程度格納できる。

【データのリード・ライト手順】

- ①必要なODをセルから選択し、アクセッサによりODにセットする。(マウント操作)
- ②リード・ライト要求を発行する。
- ③リード・ライト終了後は、ODをセルの元の位置に返却する。(デマウント操作)

【特徴】

- ・大量のデータを格納、保管できる(数十GB程度)
- ・単位容量当りの価格は磁気ディスク装置に比べ安い
- ・ランダムアクセスが可能
- ・データの書換えが可能(書換型OD使用時)

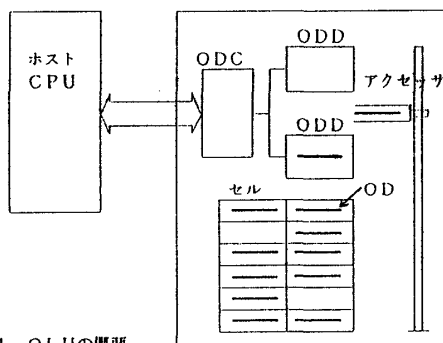


図1 OLUの概要

3 OLU利用上の問題点

本章では、DBMSにおける長大データの管理の概要を述べ、長大データの格納媒体としてOLUを使用するときの問題点を整理する。

【長大データ管理の概要】

- ①長大データの転送はデータ量が多いため、1データをAPで処理可能な大きさに分割して行う(分割転送)。
- ②長大データは制御情報とデータ実体に分離し、制御情報は文字・数値データと同じ主格納空間に、データ実体は補助格納空間に格納する(図2)。

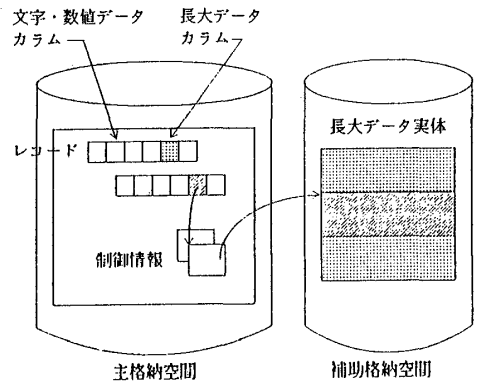


図2 長大データ管理の概要

【問題点の整理】

長大データの実体を格納する補助格納空間をOLUで構成する場合を前提に、問題点を以下に示す。

【問題1】・・・長大データのアクセス性能の問題

媒体(OD)のマウント/デマウント操作は、それぞれ十秒近い時間を要する。従って、1データの分割転送中に必要以上の媒体操作を伴うと、長大データアクセスにおけるオーバヘッドが増えアクセス性能が低下する。また、このオーバヘッドを少なくするために、1データの分割転送中、媒体をマウントし続けると、会話処理等、その間にデータの通信処理を伴う分野ではODDの保留時間が問題となる。

【問題2】・・・アクセス競合の問題

同時にアクセスできる媒体はODD数で制限される。このため、①アクセス要求がOLU内の媒体に対しランダムに発生する場合はODDの競合が発生しやすい、②複数のODDを要求する処理(コピー等)ではODDを介したデッドロックが発生する、等の問題がある。

【問題3】・・・ロールバック処理の問題

媒体が非常駐のため、ロールバック処理では媒体を確保する必要がある。特に、更新した長大データが多数の媒体に渡るときは全ての媒体をロールバックする必要があり、処理に時間を要する。また、ロールバック処理時にODDを使用するため、通常の長大データアクセス処理の性能低下を招く。

このように、従来の長大データ管理方式をそのまま適用すると性能等の問題があり、データベース格納装置としてのOLUの利用分野が限定される。

#### 4 OLUサポート方式

3章で述べた問題点を解決するOLUのサポート方式について以下に提案する。

##### 4.1 ワークファイル方式

【問題1、2の①の対策】

長大データの転送処理で、アクセス性能の向上、アクセス競合の緩和を図るには、媒体のマウント/デマウント操作のオーバーヘッド及び、ODDの保留時間を少なくし、OLUへのアクセス効率を高めなければならない。このためには、OLUに対するデータ転送は一括して行う必要がある（一括転送）。

【問題2の②の対策】

ODDを介するデッドロックを回避するためには、OLUに対するデータ転送は1データ毎に行い、転送終了後は直ちにODDを解放する必要がある。

【提案方式】

長大データの一括転送を可能とするため、ワークファイルを設け、一旦ここに蓄積する方式とする（ワークファイル方式）。

デッドロックを回避するため、一括転送では1長大データ毎にODD確保→データ転送→ODD解放を繰り返す方式とする。方式の概要を図3に示す。

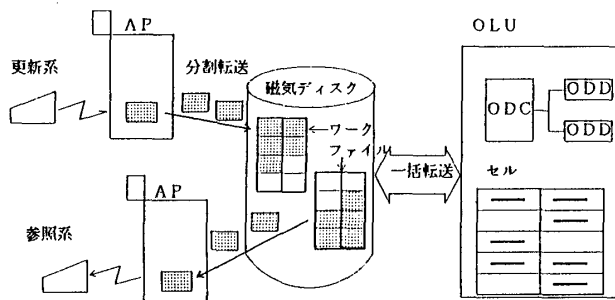


図3 ワークファイル方式

本方式では、長大データの検索/格納（更新）処理は次のように行う。

【検索処理】

検索対象の長大データをOLUからワークファイルに一括転送する。その後、APの転送要求に従ってワークファイルから長大データを分割してAPに返却する。

【格納処理】

APから分割転送される長大データをワークファイルに蓄積する。1長大データ分の転送終了を契機に、ワークファイルからOLUに一括転送する。

##### 4.2 ロールバック処理方式

【問題3の対策】

ロールバック時にODD及び、媒体の確保を不要とするためには、トランザクションの終了まで更新前データを媒体上に残す必要がある。

【提案方式】

トランザクションの終了まで更新前データを残すため、格納/更新処理では新データを旧データとは別の領域（新規領域または、そのトランザクション以外で削除されコミット済みとなった領域）に格納し、対応する制御情報を追加/修正する方式とする。

ロールバック処理では、制御情報のみをトランザクション直前の状態に書き戻す方式とする（図4）。

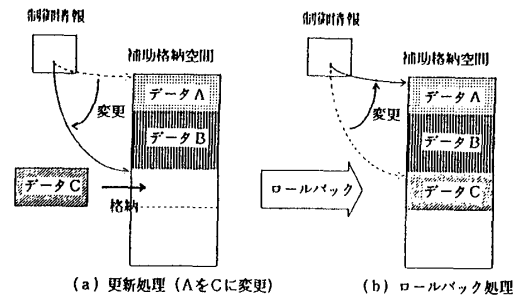


図4 長大データのロールバック方式

本方式では、迅速なロールバックが可能となるだけでなく、ロールバックに必要な長大データ実体の更新前イメージの取得も不要となる長所がある<sup>3)</sup>。

##### 5 OLUサポート方式の評価

提案方式（図3のモデル）について待ち行列モデルを適用し、そのアクセス性能を評価する。簡単のため、以下の仮定を設けOLUはM/M/2、磁気ディスクはM/M/1で近似する。

【仮定】

- ①長大データのアクセス要求の到着、各アクセスが要求する媒体はランダムとする
- ②長大データの長さの分布はランダムとする
- ③ワークファイルの容量と本数は充分にあるとする
- ④通信処理における待ち行列は無視する
- ⑤検索処理のみとする

図5に提案方式（ワークファイル方式）と、1長大データの転送中ODDを占有する方式（占有方式）の性能比較を、転送する長大データの平均長をパラメータとしたトランザクションと応答時間の関係で示す。ここで、応答時間とは端末への長大データ転送が終了するまでの時間である。図5より、性能の大幅な改善が図れることがわかる。

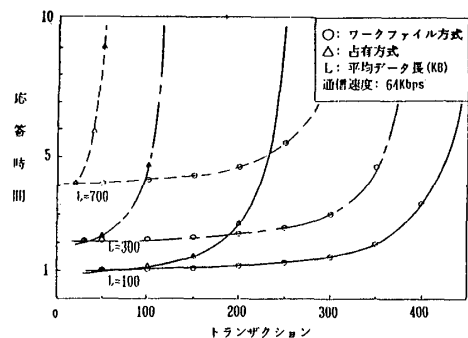


図5 OLUサポート方式の性能比較

##### 6 おわりに

長大データの格納装置として、OLUをデータベース管理システムでサポートする方式を提案した。今後は、試作システムの構築、性能の実測を行い、ワークファイルの容量と本数、OLU構成（ODD台数と格納OD数）等が性能に及ぼす影響を評価する。

【参考文献】

- 1) 芳西、長濱：「長大データの格納制御方式」、情処学会第32回全国大会
- 2) 岸本等：「マルチメディアDBMSにおける長大データ管理について」、情処学会DB研究会 1986.1
- 3) Haskin R.L., Lorie R.A.: "On Extending the functions of a Relational Database system" ACM SIGMOD, 1982