

MIAにおけるCDBからRDBのデータ移行方法

4K-5

黒岩 淳一、飯塚 哲也

NTT情報システム本部

1. はじめに

NTTでは業務システムのMIA環境への移行を推進している。それに伴い、既存業務システムで多く採用されているCODASYL型DB(以下CDBと記す)からMIAで標準とされるリレーショナル型DB(以下RDBと記す)へデータベース・データを移行する作業が必要となる。CDBの検索の特徴は、目的レコードを取り出すのみでなく、そこへ到達するまでのアクセス・パスや順序も重要なことである。このため、データベース・データ(以下レコード・オカレンスと記す)に加え、レコード・タイプ間(オナ/メム間)の親子関係(以下セット関係と記す)や同一レコード・タイプ内の順序関係(以下順序関係と記す)も移行を要求される場合がある。従って、業務目的を考慮した上でのRDB表現(セットと順序関係を考慮)方法、さらに移植後の性能劣化を最小限に留めるためのDB設計等、全体としてバランスのとれたDB設計・移行方法を心がけることが重要となる。本稿では、レコード・オカレンスをRDBへ移行する際の、RDB表現方法、およびそれらを選択するための判断基準について示す。

2. CDB移行の考え方

CDBには、格納構造の違いから階層型とネットワーク型の2種の形式が存在する。両者の基本的差異は、図1に示すようにセット関係が1:n(階層型)、m:n(ネットワーク型)であることである。

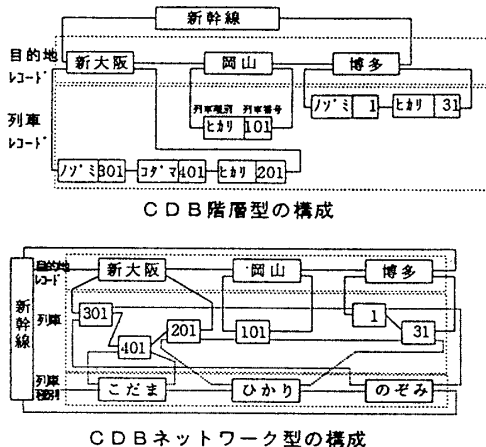


図1 CDBの階層型とネットワーク型の構造的差異

ここで、CDBにおけるデータ・アクセスについて整理すると、CDBの検索では目的レコードの取り出しに加え、該レコードに到達するまでのアクセス・パスやアクセス順序も重要である場合が多い。すなわち、レコード・オカレンスに加え、レコード・タイプ間のセット関係や同一メンバーレコード・タイプ内の順序関係もCDBアクセスにおける重要な情報の一つである。従って、CDBからRDBへの移行を実現するに当たり、上記の情報を余すところ無くRDBシステムへ移行することも必要となる。また、既存システムでのDBアクセス法(CDB重視)と移行後のRDB形態(RDB

重視)のどちらかを重視するかでDB設計の方向性が異なるため、配慮が必要となる。

ここでCDBのレコード・オカレンスをRDBに移行する観点から両者の構造的差異を考えてみると、オナを識別するための情報量に差異が生じるのみで、移行方法自体に両者の違いはないと考えることが出来る。図1に示したCDBの各々の構造において、「列車レコード」を主としてRDBに移植する際の一例について図2に示す。なお、レコード・タイプ間のセット関係を表現するために上位(オナ)レコードを特定する「上位識別子(例えば、上位レコード・タイプ'のキー値)」、オナ配下の同一メンバー内の順序関係を示すために「自識別子(例えば、順番)」を各々付加した例を示した。

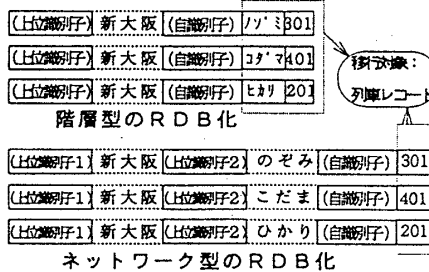
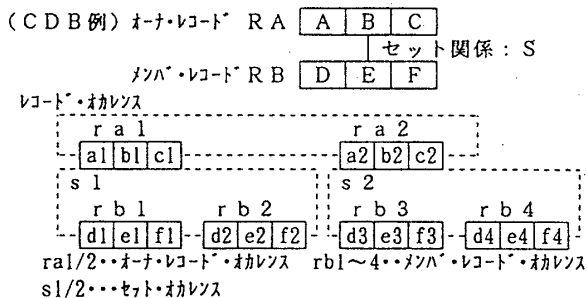


図2 CDBデータのRDB化(例)

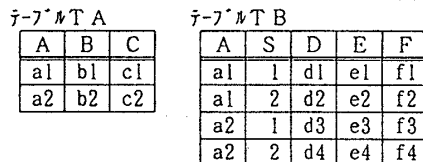
3. RDBデータへの移行方法

前節にて、オナとメンバーの1レコードを一つのペアとして表現し、RDBへ移行した例を示した。このセット関係を含めてRDBへ移行する方法は、DBシステムの用途に応じて以下の3つの型(5つのパターン)が考えられる。なお、下記の「CDB例」に基づき、各型/パターンを紹介する。



・RDBイメージ重視型

①オナとメンバーを各々1テーブルとして表現



- ・RA、RBを別々のテーブル(TA、TB)で表現
- ・TBにTAとのセット関係を表す上位識別子(A)と順序関係を表す自識別子(S)を付与

②オーナーと配下の1メンバを1テーブルで表現 (前節の例と同様)

A	B	C	S	D	E	F
a1	b1	c1	1	d1	e1	f1
a1	b1	c1	2	d2	e2	f2
a2	b2	c2	1	d3	e3	f3
a2	b2	c2	2	d4	e4	f4

- ・R AとR Bを統合して1テーブル(T A)で表現
- ・R Bの1ワード・カラムとそれに対応するオーナーAのワード・カラムのペアをT Aの1行で表現
- ・T AにR Bのワード・カラム内の順序を表す自識別子(S)を付与

・CDBイメージ重視型

③オーナーと同一タイプの全メンバの2テーブルで表現

テーブルT A			テーブルT B						
A	B	C	A	D	E	F	D'	E'	F'
a1	b1	c1	a1	d1	e1	f1	d2	e2	f2
a2	b2	c2	a2	d3	e3	f3	d4	e4	f4

← R B → ← R B →

- ・R A、R Bを別々のテーブル(T A、T B)で表現
- ・R Bのワード・カラム内の全メンバ・ワードをT Bの1行で表現
- ・T BにT Aの上位識別子(A)を付与

④オーナーと同一タイプの全メンバを1テーブルで表現

A	B	C	D	E	F	D'	E'	F'
a1	b1	c1	d1	e1	f1	d2	e2	f2
a2	b2	c2	d3	e3	f3	d4	e4	f4

- ・R AとR Bを統合して1テーブル(T A)で表現
- ・R Bのワード・カラム内の全ワード・カラムとそれに対応するオーナーAのワード・カラムのペアをT Aの1行で表現

・拡張性重視型

⑤セット関係を別にして3テーブルで表現

テーブルT A			テーブルT S		テーブルT B			
A	B	C	A	S	D	D	E	F
a1	b1	c1	a1	1	d1	d1	e1	f1
a2	b2	c2	a1	2	d2	d2	e2	f2
			a2	1	d3	d3	e3	f3
			a2	2	d4	d4	e4	f4

- ・R AとR Bを別々のテーブル(T A、T B)で表現し、セット関係を管理する別テーブルT Sを設置
- ・T SにはT A/T Bの識別子(カラムA/D)を設け、R Bのワード・カラム内の順序を表す識別子(S)を付与

4. 移行方法の評価

CDBをRDBに移行する方法として3つの型(5つのパターン)が存在することを前述した。しかし、移行を実施する際にはDBの形態(RDB/CDBイメージ)、データ量(ファイル容量)、DBアクセス性能(ディスクのI/O回数、JOINの回数等)やAPプログラムの記述量等を総合的に判断してどの方法を用いるかを選択する必要がある。

表1に各移行の型/パターンの評価方法について整理した結果を示す。

ここで、各パターンを決定する際のポイントを以下に示す。

(1) RDBイメージ重視型(パターン①または②)

- (a)セット関係をRDBで表現した場合に正規形となるパターン①が適当。但し、オーナーからメンバを検索する際にJOINが必要となるため、このような検索の必要性が少ないときや移行先DBMSでのJOIN性能に問題がないことが条件。
- (b)パターン①に問題があるならばパターン②。但し、オーナー部分のデータを重複して持つため、データ量が多くなる(5パターン中で最も大)。

表1 移行の型/パターンの評価

分類	評価項目	RDB重視		CDB重視		拡張
		①	②	③	④	
1	テーブル数	2	1	2	1	3
2	レコード長	短	中	長	長	短
3	NULLデータの有無	なし	なし	あり	あり	なし
4	最大メンバ数を考慮したカラム設計	不要	不要	要	要	不要
5	ワード間で重複した値を持つカラム	あり	あり	なし	なし	あり
6	付加するカラムとその種類	キー・順序	順序	なし	なし	キー・順序
7	D生成ワード数	1+n	n	2	1	1+2n
8	量総データ量	中	最多	小	最小	多
9	性オーナー検索時の返却ワード数	1	n	1	1	1
10	能メンバ検索時のJOINの有無	あり	なし	あり	なし	あり
11	メンバ検索時の返却ワード数	n	n	1	1	n
12	APメンバの抽出処理要否	不要	不要	要	要	不要
13	記述ワード間でオーナー部分の重複値の排除	不要	要	不要	不要	不要
14	量メンバへの条件指定	可能	可能	難	難	可能

*1: キー...上位レコード・タイプのキー、上位識別子
 *2: メンバ・レコードの数をnとする
 (注) ○...他パターンとの比較で相対的に優れている
 △... // 中間にある
 ×... // 劣っている

(2) CDBイメージ重視型(パターン③または④)

- (a)1回のSQL命令でオーナー配下の全メンバを検索可能。
- (b)RDB重視型に比べカラム数が多い、メンバの最大数分のカラムを固定的に定義しておく必要がある。
- (c)パターン③、④の差異はJOINの有無(④ではJOIN要)。

(3) 拡張性重視型(パターン⑤)

- (a)セット関係の情報を独立した別テーブルで表現するため、該情報を修正する場合の影響範囲が小さい。
- (b)反面、オーナーからメンバを検索する場合には、JOINが2回必要となり性能的には不利。

上記の様、移行後の要求業務、性能等の種々の条件を考慮し、システム設計者側にて最良の移行方法を決定する必要がある。

5. おわりに

本稿では、CDBからRDBへデータベース・データ(オカレンス)を移行する際の基本的考え方を示し、さらに考慮が必要となるRDB表現方法の型/パターンを列記し、システム設計者にとって最良なパターンを選択・決定するために必要な評価項目、判断ポイントについて整理した。

【参考文献】

- [1]水嶋宏也 他「社内システムへのMIAの導入促進(2)」N T T技術ジャーナル, Vol. 4, No. 5, pp. 17-19, 1992