

5 R-3

河村 元夫¹, 佐藤 裕幸², 横田 一正¹

1:(財)新世代コンピュータ技術開発機構

2:三菱電機(株)情報電子研究所

1 はじめに

第五世代コンピュータプロジェクトでは、知識情報処理システム(KIPS)の中核的機能であるデータベース、知識ベース管理機能を提供することを目的とし知識ベース管理システム(KBMS)の研究開発を行ってきた[1]。そのなかで、Kappa(Knowledge Application-oriented Advanced Database and Knowledge Base Management System)[2, 3]は、KIPSやKBMSのデータベースエンジンの役割を果たし、知識表現言語/知識ベース言語 *QUICKOTE*はその上位層にあたる。中期に逐次推論マシンPSIで動作する逐次データベース管理システム(DBMS) Kappa-IIを開発し、後期には並列推論マシンとそのオペレーティングシステムPIM/PIMOSの環境で動作する並列DBMS Kappa-Pの試作をおこない、現在初版が動作している。本稿では、Kappa-Pの設計方針と特徴を述べる。

2 設計方針

KIPSが扱うデータや知識は複雑な構造をしており、かつ量も膨大なものがある。たとえば、遺伝子情報処理の分子生物学データベース GenBank/HGIR database[4]は、塩基配列と特徴の記述と関連する文献情報からなり、この塩基配列は短いものから非常に長いものまでさまざまである。このような複雑なデータを伝統的な関係モデルで扱うには、データ表現能力の点と効率的な問合せ処理の点で問題が多い。また、このデータベースは、解析技術の進歩にもない近年飛躍的に増大している。さらに、このデータに対する操作は、類似検索など計算能力を必要とするものが多く、そのデータ量と計算量から、並列マシンの能力が必要になっている。

Kappa-Pが動作する環境は、並列推論マシンPIMとそのオペレーティングシステムPIMOSの環境である。PIMはMIMD型の疎結合と密結合が混合されたハイブリッド並列マシンある。10台程度の要素プロセッサが、共有バス/共有メモリにより結合され一つのクラスタとなり、それが、ネットワークで結ばれている。共有メモリは、数百メガバイトで、ディスクは各クラスタに取り付けることができる。このような背景のもと、つぎの設計方針で設計された。

- データモデル
複雑な構造データを扱えること
大量データに対する考慮
- ハードウェア資源の有効利用
疎結合並列処理、密結合並列処理の両方を考慮
大容量主記憶を生かした主記憶データベース機能
ディスクに対する並列アクセス
- 並列マシン上のDBMS
応用プログラムとの通信量の削減

3 特徴

前節の方針に基づき、PSI上で実装された自然言語処理システムや遺伝子情報処理システムで有効利用されている Kappa-IIの経験を最大限に生かして、Kappa-Pを設計した。つぎのような特徴がある。

• 非正規関係モデル

複雑な構造データを効率的に扱うために非正規関係モデルを採用する。これは、関係モデルの自然な拡張になっており、*QUICKOTE*の無限構造をもたないオブジェクト項のクラスに対応する。

データ構造としては、属性が階層構造を持ち、値として繰り返し値が許されることが、関係モデルとの違いである。また、その意味論として、非正規関係に固有の操作である行ネスト操作に対し独立した意味を与えている。つまり、非正規関係 $\{[a/c_1, b/\{c_2, c_3\}], [a/c_1, b/\{c_3, c_4\}]\}$ と $\{[a/c_1, b/\{c_2, c_3, c_4\}]\}$ は同じ意味を持っている。これにより、ネスト構造を意識しない問合せが可能となる。先の非正規関係に対し問合せ $?-[a/X, b/\{c_2, c_4\}]$ を出すと、どちらも $X = c_1$ が結果として得られる。これは、意味論としては基本的に関係に基づきながら表現(および蓄積構造)の効率化をめざしているという点で、関係モデルの自然な拡張であり、多値従属性をもつデータの表現と処理の効率化が可能である。この意味を反映して、関係代数を、行、列のネスト、アンネスト操作を含む拡張関係代数として定義しなおした。

また、さまざまな知識を格納できるようにデータ型としてタームを追加し、大量データに対する検索処理の効率化を目的に、索引としてのみ存在する属性なども追加した。

• 密結合並列と疎結合並列を考慮した構成

DBMSは大量のデータを扱うので、疎結合並列処理における通信量が大きな意味を持つてくる。そのため、密接に関連するデータを同じクラスタに置くなどのデータの配置が重要で、さらに、問合せ処理において通信量が少なくなるようなプランを立てる必要もある。これらをおこなうためには、分散データベースの構成が向いている。

図1に、Kappa-Pの全体構成を示す。各クラスタにローカルDBMS(LDBMS)と呼ばれるDBMSを配置し、全体で一つのデータベースを管理する。このLDBMSは、それ自身でDBMSとしての全機能を持ち、複数LDBMSが関与する問合せを処理するために、二相コミットプロトコルに基づく分散トランザクションの機能を持っている[7]。また、複数のLDBMSにより一つのデータベースを管理するためテーブル名などの大域情報の管理が問題になるが、それを管理するサーバDBMS(SBDM)の複製を作ることにより、アクセスの集中を回避する[6]。クラスタに割り当てられた、ローカルDBMSは、その内部処理で密結合向きの並列処理をおこなう。

問合せ処理は、インタフェースプロセス[5]と呼ばれるプロセスを介して行われる。アプリケーションが並列言語で記

Overview of the Parallel Database Management System: Kappa-P
Moto KAWAMURA¹, Hiroyuki SATO², Kazumasa YOKOTA¹
1: Institute for New Generation Computer Technology
2: Mitsubishi Electric Corporation.

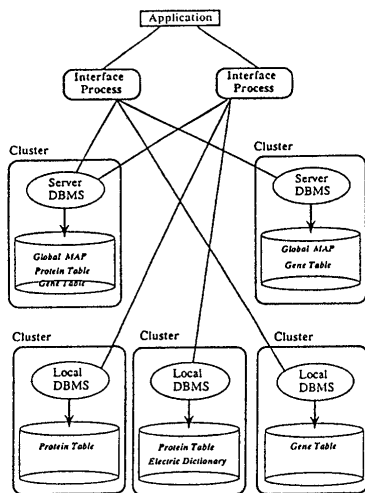


図 1: Kappa-P の全体構成

述されるため、一つのアプリケーションが複数のインタフェースプロセスを利用できるようにしている。また、このインタフェースプロセスが、問合せ処理時に必要となる各種の変換をおこない該当 LDBMS に対する部分問合せに変換する。

• データの配置

データの配置は、並列処理に関する事柄である。

クラスタ間並列処理をおこなうためには、複数の LDBMS にデータを分散配置する必要がある。これには、単純にテーブルを分散配置すること、テーブルの水平分割、テーブルの複製などがある。テーブルの水平分割は、一つのクラスタでは扱い切れないような大量のデータや特にデータ検索の速度を重視する場合などに有効である。しかし、すべての問合せに有効ではなく、逆に遅くなる場合もあるので注意しなければならない。また、テーブルの複製に関しては、現在の実装では、サーバ DBMS の大域情報管理でのみ実現されている。

クラスタに割り当てられる LDBMS 内でデータを主記憶に置くか二次記憶に置くかの選択ができる。これは、PIM の各クラスタがもつ数百メガバイトの大容量主記憶を有効利用するための機能である。主記憶上のテーブルは、二次記憶とは無関係な一時テーブルであるが、多数の中間結果テーブルを生成するシステム、たとえば演繹データベースなどでは、これが有効である。

この主記憶上の一時テーブルを利用して、擬似主記憶データベースの機能も提供する。これは、二次記憶テーブルの複製として主記憶上の一時テーブルを作り、読み系操作は一時テーブルに対しおこない、更新系操作は両方のテーブルに対して行う。これにより、更新の二次記憶への反映を保証する主記憶データベースに近い効果をねらっている。更新時は、両テーブルに対するアクセスになるが、これにはクラスタ内並列処理が有効に働く。

• 問合せ処理

Kappa-P は原始コマンドと KQL の二種類のコマンドを提供している。

原始コマンドは、非正規関係のためのプリミティブな操作を提供している。それは、レコード識別子によるポインタ操作を基本にした操作で、KQL に比べ低レベルではあるが効率的な操作が可能である。非正規関係は属性値に繰り返し許されるため、このレコード識別子も単純な構造ではなく、このレコード識別子に対応するレコードを得るにはアンネスト/ネスト操作 [9] が必要になる。

KQL は、拡張関係代数からなる式の集まりからなり、問合せ中に一時的な操作を定義したり、推移閉包を処理するためのループを記述したりできる。この KQL は、一つのトランザクション単位でまとめて受けとられ、各種最適化され、拡張関係代数処理のための中間言語 [8] への変換され、通信量を考慮し各ローカル DBMS への部分問合せに分割され、各ローカル DBMS に送られ実行される。

• Kappa-II との互換性

PIM は、そのフロントエンドプロセッサとして PSI を使っている。その PSI 上の Kappa-II 用として開発されたプログラムを利用できるようにするため、Kappa-II と互換性をもつプログラムインタフェースも提供している。これは、並列処理向けのストリームを基本にしたインタフェースを、逐次言語向きにしたものである [10]。

4 まとめ

Kappa-P は現在初版が PIM/m 上で動作しており、蛋白質データベースを格納し、そのデータベース用に Kappa-II 上で開発されたアプリケーションが動作している。また、水平分割テーブルによる並列効果も得られ、十分な性能で動作することがわかった。今後、計測結果に基づく評価、システムの改良と安定化、複数 LDBMS を利用した問合せ処理実験などをおこなう予定である。

参考文献

- [1] K. Yokota and H. Yasukawa, "Towards an Integrated Knowledge-Base Management System — Overview of R&D on Databases and Knowledge-Bases in the FGCS Project", *FGCS'92*, 1992.
- [2] K. Yokota, M. Kawamura and A. Kanaegami, "Overview of the Knowledge Base Management System (Kappa)", *FGCS'88*, 1988.
- [3] M. Kawamura, H. Sato, K. Naganuma, and K. Yokota, "Parallel Database Management System: Kappa-P", *FGCS'92*, 1992.
- [4] "GenBank/HGIR Technical Manual", *LA-UR 88-3038*, Group T-10, MS-K710, Los Alamos National Laboratory, 1988.
- [5] 永沼ほか, "Kappa-P の並列問い合わせ処理", 第 45 回情報学全国大会, 5R-04, 1992.
- [6] 坂下ほか, "Kappa-P のネームサーバ機能", 第 45 回情報学全国大会, 5R-05, 1992.
- [7] 梶野ほか, "Kappa-P のトランザクション制御", 第 45 回情報学全国大会, 5R-06, 1992.
- [8] 澤部ほか, "Kappa-P の中間言語処理", 第 45 回情報学全国大会, 5R-07, 1992.
- [9] 川村ほか, "Kappa-P のアンネスト/ネスト処理", 第 45 回情報学全国大会, 5R-08, 1992.
- [10] 中嶋ほか, "Kappa-P の単一レコード・アクセス機能", 第 45 回情報学全国大会, 5R-09, 1992.