

データベースプロセッサ R I N D A の
アーキテクチャ

5Q-4

井上 潮、 速水治夫、 福岡秀樹、 鈴木健司
NTT情報通信処理研究所

1 はじめに

リレーショナルデータベース (RDB) 処理性能の大幅向上を目的として、RDB処理の中で重要性が高く、汎用CPUにとって負担が重いサーチ処理、ソート処理を超高速に実行するデータベースプロセッサ R I N D A を開発した。本稿では、R I N D A の開発背景、アーキテクチャ、効果について述べる。

2 開発の背景

(1) 従来の問題点

従来のRDBでは、予めインデックスが作成されていない表の列を指定した検索処理を行おうとすると、結果が得られるまでに多大の処理時間とCPU負荷がかかった。このため、実際にはユーザはRDBの機能の限られた範囲内でしか利用できず、以下のような問題を抱えていた。

- ① 使用できる検索条件に融通性がなく、固定的なパターンでのDBアクセスしかできない。
- ② 大容量の表を対象とした統計処理には時間がかかり、DBから最新の情報を引き出せない。
- ③ DBの更新時にインデックスを保守する処理がシステムの性能ネックの要因になる。
- ④ 文章情報をDB化する場合、インデックスを作成するためのキーワード抽出に膨大な工数がかかる。

(2) 開発目的

上記の問題を解決するために開発した R I N D A は、汎用CPUにとって負担が重い処理を専用ハードウェアによって超高速に実行することにより、検索処理時間の大幅短縮とCPU負荷の大幅削減を実現するRDB処理専用のプロセッサである。具体的には、①に対しては複雑な条件式でのサーチ処理と結合処理の高速化、②に対してはソート処理と計数処理の高速化、③と④に対してはテキストサーチを含む高速サーチ機能による必要インデックス数の削減を可能とする。

(3) 設計条件

システムの柔軟性、拡張性を確保するとともに、既存システムへの導入を可能とするため、以下の条件で R I N D A のハードウェアおよびソフトウェアを設計した。

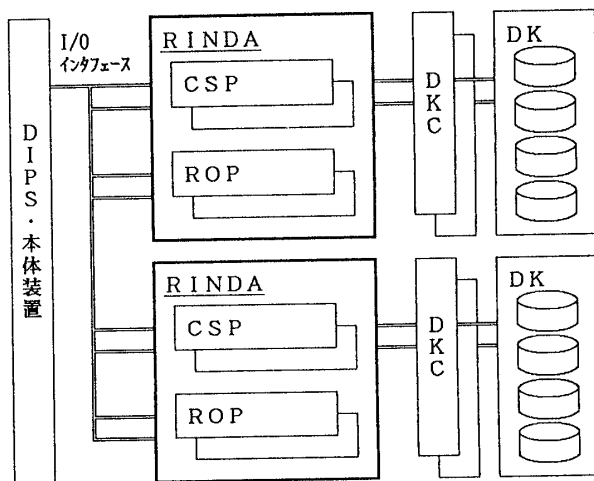
- ① 本体装置とはI/Oインタフェースで接続することによりDIPSシリーズの各機種と接続可能とする。
- ② 専用ハードウェアをコンポーネント化することにより、任意の台数をシステムに接続可能とする。
- ③ SQLインタフェースの提供により、業務プログラムからは専用ハードウェアの意識不要とする。
- ④ 1つのDBに対して、従来のインデックスを使用したアクセスとの混在実行、同時処理を可能とする。

3 ハードウェア構成

(1) 構成概要

R I N D A は、CSP (内容検索プロセッサ) とROP (関係演算プロセッサ) により構成され、それぞれ独立のI/Oインタフェースで本体装置と接続される (図1)。

CSPは、ディスクに格納された表を指定された条件でサーチし、条件に合致した行の中から処理に必要な列のみを取り出して本体装置へ転送する。ROPは、本体装置から転送されてきた表を指定された条件でふるい落としおよびソートを行い、その結果を本体装置へ返送する。CSPとROPの処理はいずれもチャンネルまたはディスクのデータ転送速度のパイプラインで行われるため、データ転送以外の処理時間はほとんど必要としない。



CSP : 内容検索プロセッサ、 DKC : ディスク制御装置
ROP : 関係演算プロセッサ、 DK : ディスク装置

図1 システムの論理構成

(2) 処理方式

ユーザがSQLで記述した問い合わせは、本体装置上のDBMSの制御のもとに基本的には、①CSPによるサーチ処理、②ROPによるソート処理、③DBMSによる出力処理の順で実行される(図2)。

なおCSPとROPをハードウェア的に独立化したことにより、複数のディスクに分散して格納された表を1台のROPでまとめてソートすることが可能である。

4 ソフトウェア構成

(1) 構成概要

DBMSソフトウェアは、SQLの解析と最適化および実行制御を行う言語処理部と、基本的なデータ入出力と資源管理を行うアクセス制御部により構成される。DBMSの従来機能を保証しつつRINDAによる高速検索機能を実現するため、言語処理部に専用の最適化機能を追加するとともに、メモリ管理、排他制御、救済制御等の一元管理が必要な部分を除き、アクセス制御部は既存のものと同様にRINDA専用のもので二本立ての構成をとった(図3)。

(2) 特有な機能

DBMSの機能のうちRINDA対応の特有なものを以下に示す。

- ① 全ての処理を複数行単位で行うことにより、CSPとROPの処理効率を向上させる。
- ② CSPの出力データ量をもとにした動的な最適化を行うことにより、以後の処理時間を短縮する。
- ③ CSP、ROPの連動制御を行うことにより、処理に必要な本体装置のメモリ量を削減する。
- ④ 複数CSPでの並列サーチ処理を行うことにより、大容量の表の検索処理時間を大幅に短縮する。

5 効果

(1) CSPの効果

表全体のサーチをインデックスを使用せずに行う場合、従来の処理方式では1ページ単位でディスクへのI/Oを起動し、1行単位でCPUによる検索条件評価を行う。一方RINDAの処理方式では、ディスクからの複数ページ連続読み出し、検索条件評価、条件に合致した行の転送を全てCSPが実行する。このため、ディスクのシーク・サーチ回数削減によるI/O時間の大幅短縮とCPU負荷の大幅削減が達成できる。

(2) ROPの効果

大量の表のソートを行う場合、従来の処理方式ではCPUによる1行単位の比較処理による並べ替えを行うとともに、中間結果の格納のために作業ディスクを使用する。一方RINDAの処理方式では、CPUとROP間で往復の

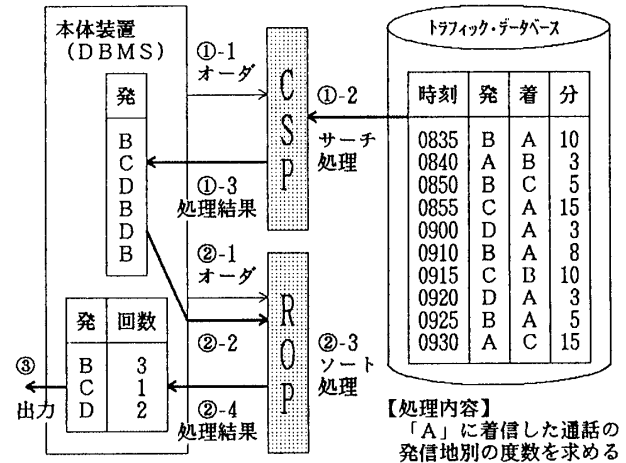


図2 処理方式

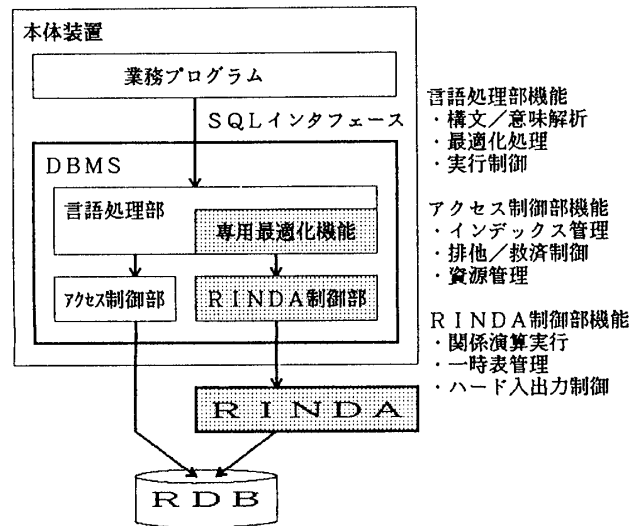


図3 DBMSの構成

データ転送を行うだけでソート処理が完結する。このためCPU負荷の大幅削減が達成できるとともに、ソート処理用の作業ディスクへのI/Oが不要になる。

6 おわりに

データベースプロセッサRINDAのハードウェア構成とこれを制御するDBMSソフトウェア構成の概要を述べた。RINDAは汎用CPUにとって負担が大きいサーチ処理、ソート処理を超高速に実行することにより、インデックスを利用できない場合の検索処理時間の大幅短縮を実現した。同一のDBに対してRINDAを使用したアクセスとインデックスを使用したアクセスを併用することにより、ユーザが利用できるRDBの機能範囲が一段と拡大して、今後さらに大容量のDBに対してもRDBが普及していくものと期待できる。