

# データベースプロセッサRINDAの 検索処理速度向上効果

黒岩淳一、板倉一郎、井上 潮、福岡秀樹  
NTT情報通信処理研究所

3Q-3

### 1. はじめに

リレーショナルデータベース処理性能の大幅向上を目的として、サーチ処理、ソート処理を高速に実行するデータベースプロセッサRINDAを開発した[1]。本稿では、インデックスを用いない選択・行計数関数・テキストサーチ、及び行計数関数とテキストサーチの組合せについて、ソフトウェアのみで検索を行った場合に対するRINDAの性能向上効果を示す。

### 2. 性能評価の目的

RINDAはインデックスが使用できない、又はインデックスの効果が無いリレーショナルデータベースの検索処理の性能向上を図るため、開発された専用プロセッサである。RINDAは、CSP(内容検索処理機構)とROP(関係演算処理機構)から構成される。CSPは、DK上の表をDKの転送速度に追従してオンザフライでサーチ処理を行い、サーチ条件に合致する行の中から必要な列だけを取り出し本体装置に転送する。

本稿では、表の全ページをサーチする検索について、CSPの性能向上効果を明らかにする。

### 3. 性能評価条件

#### (1) システム構成

評価に用いたシステム構成は、RINDAのCSP 2台、DIPS-V30 [2]、ディスクコントローラ(DKC) 2台、ディスク(DK) 2台で構成される。(図1)

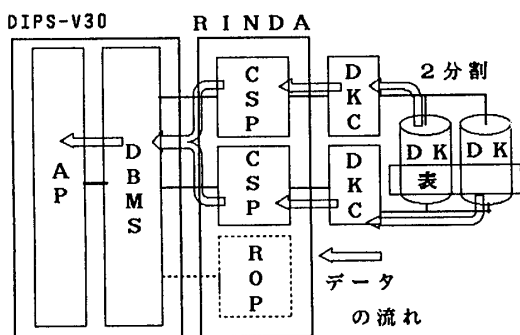


図1 評価システムの構成

#### (2) 性能評価モデル

評価モデルには、Wisconsinモデル[3]、[4]の1万件(約2MB)と10万件(約20MB)の表を利用した。各表は4B整数が13列、52B文字列が3列の計208Bの

行から構成されており、各列の値はランダム又は昇順に生成されている。なお、RINDAで性能向上効果が大きいテキストサーチ処理の評価を行うため、文字列に対する条件により、テキストサーチ処理のヒット件数を変えることが可能なように文字列を作成した。また、CSPの並列検索を可能とするため、各表は2つのDKに分割して格納した。

#### (3) 測定ステートメント

測定ステートメントを表1に示す。

表1 測定ステートメント

通番	検索種別 及び SQLステートメント例
①	選択(1件選択)
②	選択(1%選択)
③	選択(10%選択) (例) SELECT * FROM TENKA WHERE UNIQUE2=5000 AND UNIQUE2<6000
④	行計数関数(1%計数)
⑤	行計数関数(10%計数) (例) SELECT COUNT(*) FROM TENKA WHERE UNIQUE2=5000 AND UNIQUE2<6000
⑥	テキストサーチ(1%選択)
⑦	テキストサーチ(10%選択) (例) SELECT * FROM TENKA WHERE STRING1 LIKE '%AAAAAAAAAAAAAAAAAA%'
⑧	行計数+テキストサーチ(10%選択・計数) (例) SELECT COUNT(*) FROM TENKA WHERE STRING1 LIKE '%AAAAAAAAAAAAAAAAAA%'

表1に示す①~③はWisconsinモデルで定義されたステートメントであり、④~⑧はRINDAの効果が特に期待できる行計数関数、テキストサーチ性能の測定のために追加したステートメントである。

#### (4) 測定方法

性能測定の対象とする時間は、以下の通りとした。

- ・ RINDAを制御するDBMSが検索を開始してから、検索結果をDK上の一時表(検索結果を格納する作業用の表)に格納するまでの経過時間。

### 4. RINDAの検索処理速度向上効果

図2に選択、行計数関数、テキストサーチ及び行計数とテキストサーチの組合せの性能測定結果を示す。

なお、図中で用いた性能向上効果とはRINDAを使用しない場合と使用した場合の検索処理時間の比であり、また、検索率とは検索の結果ヒットする行の数と検索対象と

なる表全体の行の数との比である。

図2からRINDAの使用により30倍から140倍の性能向上が得られることが確認できる。

(1) 選択

RINDAの使用により30倍から50倍の性能向上効果が得られる。

(2) 行計数

行計数処理では、性能向上効果はほぼ60倍で一定している。

(3) テキストサーチ

テキストサーチ処理の、性能向上効果は選択処理に比べて大幅に高く、80倍から130倍となる。

(4) 行計数とテキストサーチ

行計数とテキストサーチの組合せでは、130倍の性能向上効果が得られる。

RINDAを使用した検索の性能向上要因を以下に挙げる。

① CSPではサーチ処理をデータ転送速度で実行するため、検索時間とCPU時間が短縮される。

② CSPではDK上のデータをシリンダ単位で連続読み出し出来るため、I/O時間が短縮される。

ただし、検索結果を一時表に格納するため、本体装置のバッファとDKとのI/Oが検索率に比例して増加する。

なお、参考までにRINDAを基準とした場合の他のデ

ータベースマシン (DBM) との検索処理時間比を図3に示す。

5. おわりに

データベースプロセッサRINDAの選択・行計数関数・テキストサーチ及び行計数とテキストサーチの組合せの性能評価を行なった。その結果、インデックスを用いない場合のRINDAの性能向上効果はきわめて大きなものであることが確認できた。

今回はCSPの性能向上効果について評価を行ったが、今後ROPを用いたソート処理、結合処理についての性能評価を行なう予定である。

[参考文献]

- [1]井上、速水、福岡、鈴木：データベースプロセッサRINDAのアーキテクチャ、情処37全大,5Q-4,1988
- [2]伊藤、矢沢、福村、小浜：DIPS-V30の実用化、通研実報、Vol.35, No.5, 1986
- [3]Bitton, Dewitt他：“Benchmarking Database Systems A Systematic Approach”, Proc. 18th VLDB Conference, 1983
- [4]Dewitt, 他：“A Single User Evaluation of the Gamma Database Machine”, Database Machines and Knowledge Base Machines, Kluwer Academic Publishers, pp.370-386, 1988

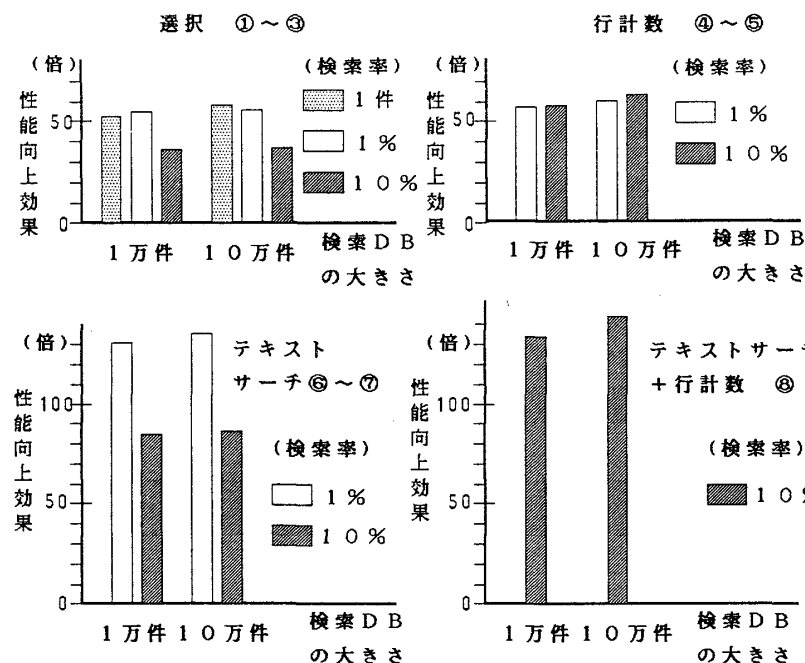


図2 RINDAの性能向上効果

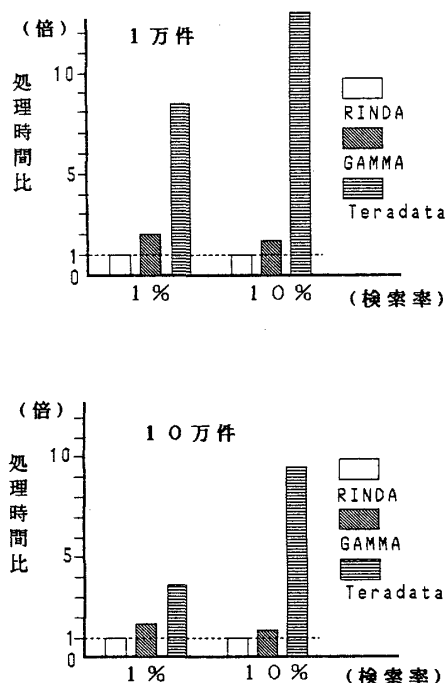


図3 他社DBMとの比較 (選択) ②~③