

高速フィルタリングプロセッサ実験システムの開発(1) 全体構想

3Q-4

北嶋 弘行、 大曾根 匡、 山本 彰
(株) 日立製作所 システム開発研究所

1. はじめに

意思決定支援システムや文書データベース検索をはじめとして、リレーショナル・データベース (RDB) を対話形式で用いる、いわゆる、非定形的処理のニーズが高い。これを高速化するための一手段として、データ転送と同期してDB検索処理を行うフィルタリング型のDBマシンの実験システムを開発した。本報告では、実験システムの狙い、方式上の方針について述べる。

2. DBマシンの目的

RDBを用いた意思決定支援システムや文書DB検索などの、いわゆる、非定形処理のニーズが高まっている。この場合、下記の理由により、インデクスのみにも頼ることは難しく、従来の計算機システムにとっては著しく高負荷となっていた。

(1) 検索条件が未知であるため、インデクスが利用できるとは限らない。しかも、過度にインデクスを設定することは、更新処理の性能劣化を招く。

(2) インデクスが付いていても、統計処理をはじめとして、インデクスでのヒット率が高ければ入出力回数が過大となり、性能劣化を招く。

(3) 文書DB検索では、インデクス(この場合は、キーワード)が有っても、キーワードの陳腐化や分解能不足の問題があり、自由キーワードによる全文検索を併用することが多い。

上記の問題を解決するために、データ転送と同期してDB検索処理を行うフィルタリング型のDBマシンの実験システムを開発した。図1にDBマシンの論理的位置づけを示す。DBマシンはチャンネルと入出力コントローラとの間に位置し、外部記憶装置からのデータ転送中とパイプラインで(いわゆるオンザフライで)データベース演算処理のうちの選択、文字列検索、射影、計数などを実行する。これにより、入出力時間を殆ど増加させることなく、CPU負荷を大幅に低減することが可能となる。

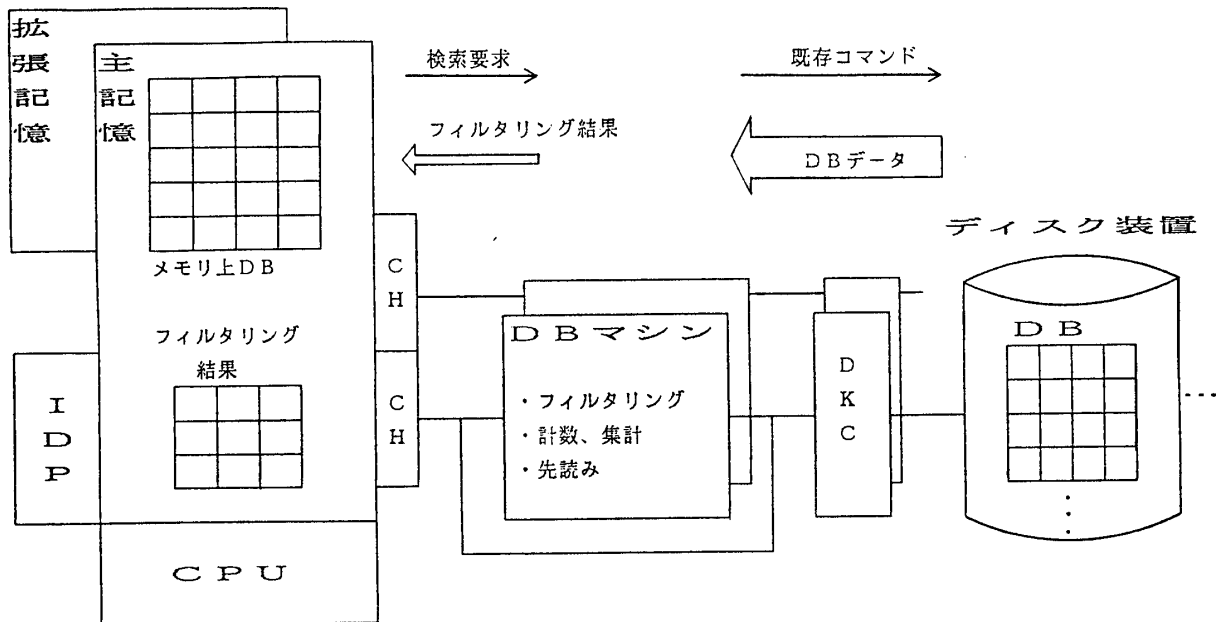


図1. DBマシンの位置づけ(1)

Experimental Filtering Processor (1) — System's Outline —

Hiroyuki KITAJIMA, Tadashi OHSONE and Akira YAMAMOTO
HITACHI, Ltd.

3. DBマシンの方式上の方針

上記のごとく、本DBマシンはインデクス無し、あるいは、インデクスによるヒット率が高い検索（以下、これらを集合検索と呼ぶ）に対するCPU負荷低減と応答性向上を目的としている。このDBマシンの適用範囲を広め、かつ、アーキテクチャとしての拡張性、連続性を高めるために、次の方針をたてた。

3.1 集合検索の高速化

(1) 専用ハードウェアによる適用範囲の拡大

複雑な検索条件式や繰り返し欄に対する検索の場合のように、高負荷な検索要求に対しても、データ転送速度に追随可能とする。しかも、データ転送速度の一層の高速化へも対応できるようにする。このため、専用ハードウェアをパイプライン動作させることとした。

(2) ソフトウェアの集合処理利用による高速化

①入出力発行、排他制御などを複数ページ分まとめて一括処理、②検索結果のDBMS部から対話処理部への引渡しの一括化。

(3) 論理的先読み方式による入出力時間の削減

DBマシンから現検索範囲の検索結果を受け取る際に、DBマシンに対し、次の検索範囲と検索条件を指定可能とする。これにより、同一入出力装置に関する、CPU処理とDBマシン処理との並列処理を可能にする。

3.2 定形処理との混在を促進

同一計算機システム内で定形処理と非定形処理の両方を稼働させる運用形態が考えられる。DBマシンは非定形処理のCPU時間を大幅に削減するので、基本的に、両者の混在を促進するが、これをさらに加速するために、主に、次の点に留意した。

(1) 入出力要求発行および排他制御の単位の分割化

定形処理との混在時には、DBマシンへの入出力要求の発行単位や排他制御の単位を数トラック以下とし、非ヒットローの排他は解放する。さらに、いわゆるデータ・リードモードを設ける。

3.3 エンドユーザ、APへの透過性

DBマシン導入が、エンドユーザの操作や運用、AP、DB構造などに、一切、影響を与えないようにする。また、DBマシンを使うか否かは、最適化機構が自動判定出来るようにする。

3.4 アーキテクチャの拡張性、連続性

(1) 各種システム構成への拡張性・・・DBマシンはチャネルと入出力インターフェイスで接続されるため、

各種のCPU、入出力装置への接続が可能である。また、集合検索の負荷の増大に対しては、DBマシンの増設で対応できる。さらに、上位クラスのCPUに対しては、複数のDBマシンを並列動作させることで効果を出す。

(2) 大容量メモリへの拡張性・・・大容量メモリ（主記憶、拡張記憶等）にデータを常駐ないしバッファリングし、システム性能を向上する方向がある。さらに、これを専用ハードウェアで更に高速化する方式としてIDPがある。しかし、少なくとも近い将来に、DB（特に、非定形処理用DB）の全てがメモリ上に載るとは考えにくい。IDPとの関係について見ると、DBマシンは2次記憶上のDBを一次フィルタリングし、IDPはこの結果や、もともとメモリ上にあるデータを対象とする。

(3) データ転送速度向上への追随性・・・2.3.1(1)で述べたように、題記を可能にしている。

4. おわりに

フィルタリング型のDBマシン実験システムの狙いと方式上の方針を述べた。DBマシンは、専用ハードウェアの追加のみで、非定形処理の大幅な高速化が可能である。また、専用ハードウェアのパイプライン化により、極めて厳しい検索条件にも十分に適用できることを狙っている。さらに、定形処理との混在環境にも適合できるだけでなく、今後の諸動向のもとで、アーキテクチャとしての拡張性、連続性に優れた方式を狙っている。

参考文献

- [1] 大曾根 他, "高速ストリング・サーチ・アルゴリズムの提案," 情報大全, pp.463-464(1987).
- [2] Torii et al., "A Database System Architecture Based on a Vector Processing Method", 3rd International Conf. of Data Engineering, 1987

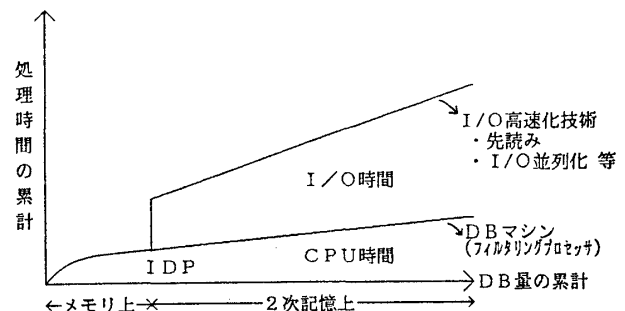


図2. DBマシンの位置づけ (2)