

SQL/GREO：データベースプロセッサGREOを用いた並列処理方式

2G-3

佐藤 重雄, 道下 学, 岩崎 浩文, 安藤 隆朗

三菱電機(株) 情報システム研究所

1 はじめに

データベースプロセッサGREO^[1]を用いたリレーショナルデータベース(Relational Database)システムSQL/GREOを構築した。本稿では、特に、SQL/GREOシステムで用いた、データベースプロセッサ使用を最適化する並列処理方式について述べる。

2 データベースプロセッサGREO

GREOは、データベース処理におけるソート処理を専用ハードウェア化(ハードウェアソート)することにより、データベース処理を中心とするビジネスアプリケーションの高速化を実現したものである。GREOのハードウェアソータは、パイプラインマージソートアルゴリズムを用いており、大容量ファイルの検索、ソート、ジョイン等を高速に実現することが可能である。

3 SQL/GREOシステムの概要

SQL/GREOシステムは、国際標準データベース言語SQLに基づく言語インタフェースを提供し、SQLの持つプログラムの容易性、データベース操作の簡易性を実現する一方、GREOを最大限に活用して、SQLで記述された複雑なデータベース処理を高速に実現するシステムである。図1に、SQL/GREOシステムの処理の概略を示す。

アプリケーション中に記述されたSQL文は、

SQL/GREO:Parallel processing with database processor GREO
Shigeo SATO, Manabu MICHISHITA,
Hirofumi IWASAKI and Takaaki ANDO
Mitsubishi Electric Corp.
5-1-1 Ofuna, Kamakura, Kanagawa 247, Japan

SQLコンパイラによって木構造のコードに変換される。木構造の各節点(node)は、GREOあるいはホストコンピュータのCPU(SQL/GREO中核部)で実行可能な処理である。SQL文で記述されたデータベース処理は、この木構造の葉レベルから、データフローの発火制御により処理を実行することによって達成される。

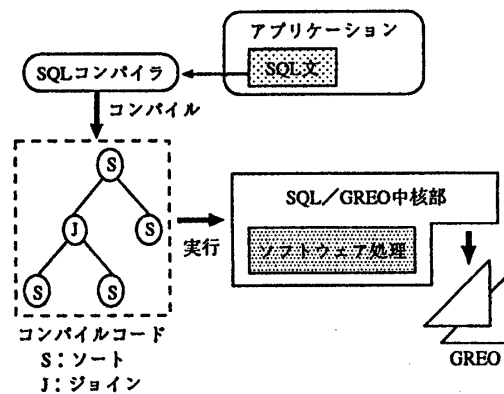


図1: SQL/GREOシステムの処理

4 GREOを指向した並列処理

4.1 GREOによる並列処理

SQLコンパイラは、GREOを有効に活用するために、ソートを主体としたオブジェクトコードを生成する。また、GREOが複数台付加されているシステムでは、上記木構造のコードに対する発火制御により、複数台のGREOで並列に処理を実行する。データベース処理において頻繁に使用されるソート処理を、高速化、並列化することにより、結合(JOIN)処理や併合(UNION)処理の高速化が可能となる。

4.2 GREOとホストCPUによる協調並列処理

GREOは、大容量ファイルのソートは高速に実行できるが、以下の場合には、その効果を十分に発揮できない。

- 対象とするファイルの容量が小さい場合。
この場合は、GREOに対する入出力のオーバーヘッドが大きくなる。
- 大容量のファイルに対し、ある列をキーにして少量の件数を抽出する場合。GREOにはファイルの全件を入力するため、キーとなる列に索引が付加されている場合には、索引を利用した処理の方が高速となる。

このような場合には、GREOを利用せず、ホストコンピュータのCPUで処理（ソフトウェアによる処理）の方が高速となる。

データベースプロセッサの使用方式としてはRINDAを用いた例^[2]があるが、SQL/GREOシステムでは、対象となるファイルの統計情報（レコード件数、索引の有無、重複値の数等）を利用して決定する。SQLコンパイラは、あらかじめ採取してある統計情報をもとに、以下のようにコード生成を行う。

- 対象とするファイルに索引が付加されている場合：索引を利用することによって絞り込んだ結果が、ある件数以下ならばソフトウェアで処理する。
- 対象とするファイルに索引が付加されていない場合：レコード件数がある件数以下ならばソフトウェアで処理する。

実行の例を図2に示す。ここで、ファイルT、Uは大容量ファイル、Vは索引を利用することによって、結果を絞り込むことができるファイルの場合である。木構造の葉の部分は並列に処理できるため、GREOとホストCPUが協調して処理を行うことができる。

この他に、ソフトウェアで処理する方が有効な場合には、GREOの負荷が高い場合、GR

EOが使用不能となっている場合がある。実行時にこのような状態を検出すると、GREO用のコードをソフトウェア用のコードに自動的に生成し直して処理を実行する。

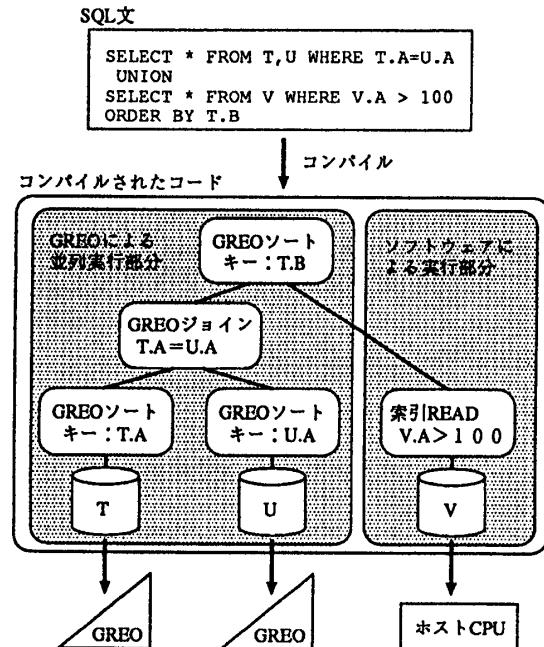


図 2: コンパイルコードの並列実行

5 おわりに

本稿では、SQL/GREOシステムの特長である高速な並列検索処理について述べた。SQL/GREOでは、検索機能の他に、データの定義機能、更新機能も備えている。GREOとソフトウェア処理との使い分けなどのチューニングを行うことにより、高速なデータベースシステムの構築が実現できると考える。

参考文献

- [1] 伏見他「データベースプロセッサGREO」、情報処理, Vol.33 No.12 1992
- [2] 芳西他「データベースプロセッサRINDAにおける問合せ処理のアクセスパス決定方式」、情報処理学会論文誌, Vol.32 No.11 1991