

並列DBサーバシステムにおける問合せ処理方式の提案

1F-2

岩田守弘 土田正士 根岸和義 正井一夫 宮崎光夫
(株)日立製作所システム開発研究所 同ソフトウェア開発本部1. はじめに

実世界においてデータベース化が望まれるデータの量は増加する一方である。多大なデータに対しデータベース管理システムはデータ量に比例しない応答時間が望まれている。そこで、データベース処理の並列化が必要とされる。並列データベース管理システムではデータを複数のプロセッサに分割し並列に処理を行ったり、各種の処理を複数のプロセッサでパイプライン実行を行うことでデータ取り出し処理や結合処理の処理時間の短縮化を図ることができる。本稿では、並列結合処理の処理時間見積りから抽出した特性に基づき、処理時間に対して最適な並列結合処理を選択する方法に関して報告する。

2. 最適な問合せ処理の選択方式の概要

問合せ言語は、どのようなデータを検索したいかの記述であり、1つの問合せに対して複数の処理手順候補が存在する。最適な処理手順を選択するうえで、次の問題点がある。C言語などのホスト言語に問合せを埋め込み使用する場合において、ホスト言語の変数が問合せに使用され、変数の値は実行時まで決定されないことである。この問題を解決するために最適化をコンパイル時と実行時の2つのフェーズに分け、コンパイル時に処理手順の候補を実行形式に展開しておき、実行時の変数値にしたがって処理手順を選択する方式が提案されている。[1][2]

データ取り出し処理の最適化については報告されているので[1]、本稿では並列結合処理の最適化について報告する。

3. 並列結合処理方式の応答時間特性

実行時に最適な並列結合方式を選択するためには、コンパイル時に決定されるパラメータと実行時に決定されるパラメータを切り分ける必要がある。そのために、並列結合処理でも絞り込んだデータを受取る場合、インデクスを利用する場合を想定して、処理時間算出式を作成し、グラフにより選択率、ロウ数、射影ロウ長などのパラメータを変化させ評価することで並列結合処理方式の特性を抽出した。

図1に本報告で評価を行った並列ネストループ結合と並列ソートマージ結合の概要を示す。

図2は、T1の選択率を変化させ、他のパラメータを固定した場合の並列結合処理時間グラフの1例である。

各並列結合処理方式の処理時間を見積もることによって次のことがわかった。

(a) 並列ネストループ結合の処理時間は、ほとんどIO時間である。そのIO時間は、選択カラムの選択率に比例する。

(b) 並列ソートマージ結合の処理時間には対数が使われているため片側表の選択率を変化させたグラフは曲線になる。しかし、並列処理を対象とする10万規模以上のロウ数の場合には曲率(曲がる度合い)は小さい。

(c) ロウ数が数倍程度に変化しても、並列ネストループ結合と並列ソートマージ結合を切り分けるためのグラフの交点の選択率は一定。

(d) 並列ソートマージ結合のグラフの曲率が小さいため、並列ネストループ結合のグラフとの交点は1つである。

(e) データ取り出し処理方式を変えた場合の選択率の交点は、結合方式にかかわらず一定値である。

処理時間の評価結果から、次の結論が得られた。

(1) 射影ロウ長などコンパイル時に決定できるパラメータはコンパイル時に計算する。

(2) コンパイル時から実行時までの間に、大幅にロウ数が変化しないならば、ロウ数はコンパイル時に決定するパラメータとする。

(3) 変数の選択率は実行時に決定する値である。並列ネストループ結合の処理時間を選択率の1次式で、並列ソートマージ結合の処理時間を選択率の低次式で近似することで、実行時に結合方式を選択するための判別式を作成する。

(4) 実行時に決定されるパラメータを選択率1つに絞れば、グラフの交点が1つであるので、連立方程式を解くことによって、並列ソートマージ結合と並列ネストループ結合を選択するためのしきい値を求める。

(5) データ取り出し処理は結合処理方式と独立に最適な方式を選択する。

4. 並列結合処理方式の選択方法

(1) 仮定

(a) コンパイル時と実行時でロウ数が大幅に変化しない

(b) 処理を実行するプロセッサのCPU性能(MIPS)とディスクのIO性能(ランダムIO時間、一括IO時間)は、コンパイル時に処理を行うプロセッサがわかっていなくても一定値として与えることができる

Proposal of Query Processing Method for Parallel DB Server System

Morihiro IWATA, Masashi TSUCHIDA, Kazuyoshi NEGISHI, Kazuo MASAI and Mitsuo MIYAZAKI
HITACHI, Ltd.

(c) ソート処理以外は、CPU負荷よりもIO負荷の方が充分に大きい

(d) 通信にかかる時間よりも結合処理時間の方が充分に大きい

(2) 並列結合処理の選択

(a) インデックスの有無などの条件から並列結合処理手順の候補を作成 (コンパイル時)

(i) 結合する表に、結合条件カラムのインデックスが存在しない場合には並列ソートマージ結合を使用する。

(ii) 片側の表に、結合条件カラムのインデックスが存在する場合には並列ネストループ結合と並列ソートマージ結合を結合方式の候補とする。

(b) 判別式の作成 (コンパイル時)

結合する表をT1,T2として、選択条件カラムの選択率をそれぞれx,yとする。

h, u, v, w は並列処理時間の見積りに使用した関数である。

並列ネストループ結合処理時間

$$= h(\text{行数, 選択率, インデックス段数, 結合行数比, IO性能, データ分割数})$$

$$= nx \text{ と置く。 (変数の場合の選択率以外コンパイル時に決定されるので定数nとする)}$$

並列ソートマージ結合の処理時間

$$= u(\text{行数, 選択率, 射影行数, CPU性能, IO性能, データ分割数})$$

$$+ v(\text{行数, 選択率, 射影行数, CPU性能, IO性能, データ分割数})$$

$$* \log(w(\text{行数, 選択率, 射影行数, IO性能, ソートウェイ数, データ分割数}))$$

選択率以外のパラメータはコンパイル時に決定される。したがって、logを低次多項式で近似した式を、s(x,y)と置く。

$s(x,y) \leq nx$ をあらかじめ解いておき判別式とする。

片側表(T1)の選択条件カラムのみ変数の場合には、T2の選択率もコンパイル時に決定される。実行時に決定するパラメータを片側表の選択率1つのみにする場合は交点は1つであることが、3.の(d)でわかっているので方程式が解ければしきい値(Cとする)が求まる。

(c) 統計情報による絞り込み (コンパイル時)

コンパイル時の統計情報から選択率の取り得る最大値と最小値を取得し、最大値がしきい値Cより小さければ並列ネストループ結合に、最小値がCよりも大きければ並列ソートマージ結合に絞れる。

(d) 実行時最適化 (実行時)

統計情報から変数値に対する選択率を取得し、選択率がしきい値Cより小さい場合は並列ネストループ結合を、Cより大きい場合には並列ソートマージ結合を選択する。

5. おわりに

提案方式は、大規模DBを対象とし、高速なCPU

を持ち、通信処理がネックにならない、並列DBシステムの結合処理に対して有効である。また、結合する表の行数が大幅に変化しないことが前提となる。本方式により、実行時に決定される変数に対して、処理時間が短い並列結合処理方式が選択される。

参考文献

[1] 土田、武藤：RDBにおける埋込み型問合せの最適化方式の提案，情報処理学会第36回全国大会，3F-3，1988。
 [2] Graefe,G.,et al：Dynamic Query Evaluation Plans, ACM SIGMOD 89.

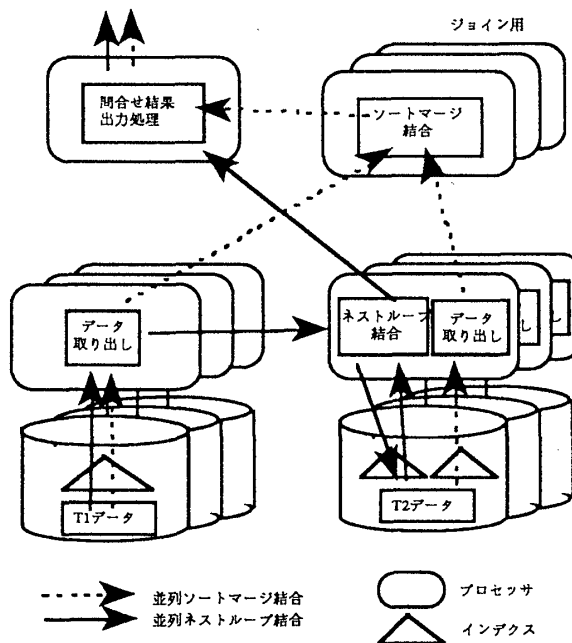


図1 並列結合処理方式の概要

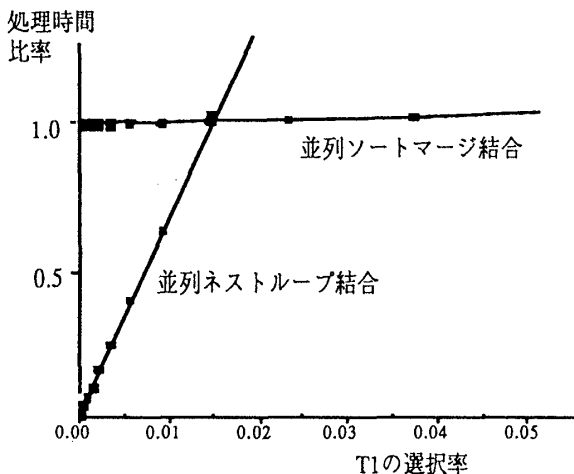


図2 並列結合処理方式の性能