

ヒストグラムを利用した SQL クエリの分割範囲計算手法の提案

齋藤 和広[†] 渡辺 泰之[†] 村松 茂樹[†] 小林 亜令[†]株式会社 KDDI 研究所[†]

1. はじめに

マルチデータベースシステム (MDBS) [1] は、データベースシステム (DBS) のクエリ処理結果に対して更にクエリ処理を継続して行う。例えば図 1 のように、MDBS が各 DBS (DB1, DB2) に対して単体で実行可能な SQL クエリを投稿し、各 DBS のクエリ処理結果に対して残りの SQL クエリ処理を MDBS 上で実行する。ここで MDBS 上のクエリ処理のデータサイズが MDBS のメモリサイズを超えると MDBS の性能が劇的に劣化し、場合によっては MDBS がエラー終了する。これは DBS のクエリ処理結果を複数回に分割して取得し、DVS 上の処理も分割実行することで回避可能である。しかしそのためには、DBS のクエリ処理の出力データサイズだけでなく、MDBS 上の処理の出力データサイズも考慮する必要がある。

そこで本稿では、MDBS のメモリサイズを超えずに SQL クエリを実行するために、ヒストグラムを利用した SQL クエリの分割範囲計算手法を提案する。提案手法は、文献 [2] で提案されている SQL クエリの出力データのヒストグラム作成手法を活用し、DBS と MDBS 双方の出力データサイズを考慮した分割範囲を計算する。また、提案手法の有効性を TPC-H クエリを用いて検証する。

2. 提案手法

提案手法では、テーブルの属性ごとに作成されるヒストグラムを用いて DBS への SQL クエリの分割範囲を計算する。ヒストグラムは各属性が持つ値の種類 (ドメイン) と、その属性内で各ドメインが出現する回数 (カウント数) で構成される。提案手法は、[2] と同様に、バケットと呼ばれる一定のドメイン範囲で集計する。バケットはドメイン範囲内に含まれるドメイン数と、それらのカウント数の合計を持つ。

MDBS が DBS に投稿する SQL クエリの分割範囲を計算するためには、①分割基準属性の選択、②分割範囲の計算の二つの手順を踏む必要がある。

分割基準属性は、分割範囲の計算においてヒ

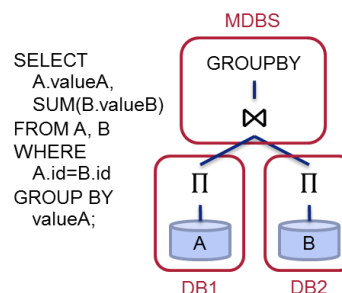


図 1 MDBS における複数 DBS 間のクエリ処理例

ストグラムを利用する属性であり、クエリの処理毎 (結合処理や集計処理等) の入出力テーブルでそれぞれ選択される。各処理の入力テーブルと出力テーブルでは同じ属性が選択され、出力されない属性は選択されない。この属性は MDBS で行われる処理に依存し、例えば結合処理における結合条件のように、処理条件となる属性が優先的に選ばれる。そのような条件がない処理では、索引付き属性や一意性制約付き属性等の制約付き属性が選択され、これらがなければランダムで選択される。

分割範囲の計算は、クエリの処理毎に、分割基準属性のヒストグラムのバケット単位で行う。分割基準属性のヒストグラムは、DBS が持つテーブルのヒストグラムを事前に集計し、[2] の方式を利用してクエリコンパイル時に作成する。まず単項演算処理の場合には、入力テーブルの分割基準属性のバケットを一つ選択し、このバケットのドメイン範囲を含む出力テーブルの分割基準属性のバケットを全て選択する。二項演算処理の場合には、上記に加えて、更にもう一方の入力テーブルの分割基準属性から、最初に選択した入力テーブルのバケットのドメイン範囲を含むバケットを全て選択する。次に、選択した全てのバケットのカウント数から利用メモリサイズを計算する。なお、選択したバケットのドメイン範囲が、最初に選択した入力テーブルのバケットのドメイン範囲外を含む場合、当該ドメイン範囲のうち範囲内となっている割合分をカウント数とする。計算した利用メモリサイズと、上限となるメモリサイズを比較し、上限を超えていなければ、同様に入力テーブルの分

Proposal of Calculation Method of Split Ranges for SQL Query based on Histograms

[†]Kazuhiro Saito, [†]Yasuyuki Watanabe, [†]Shigeki Muramatsu,

[†]Arei Kobayashi

[†]KDDI R&D Laboratories, Inc.

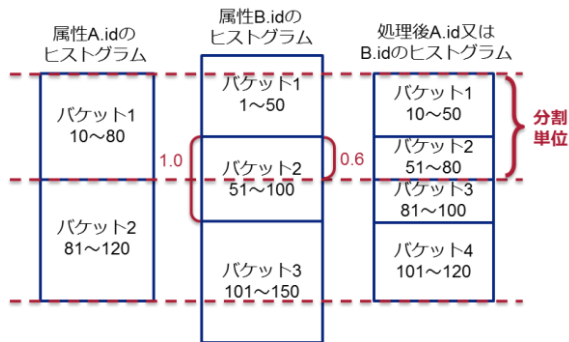


図 2 結合処理における分割範囲の計算

割基準属性の次のバケットから利用メモリサイズを計算して合算していく。上限を超えた段階で、そのバケットより前までにメモリサイズを合算してきたバケットのドメイン範囲を分割範囲として設定して、次のバケットに移行する。

図 2 は分割範囲の計算の例としてテーブル A と B の結合処理におけるヒストグラムを用いた分割範囲の計算を示す。分割基準属性 A.id のバケット 1 を基準とすると、同じドメイン範囲を含む B.id のバケット 1 及び 2 と処理後の属性のバケット 1 及び 2 の計五つを対象として、メモリサイズ計算を行う。このとき A.id のドメイン範囲を基準として、範囲外を含むバケットのカウント数を再計算する。例えば B.id のバケット 2 は、範囲内の割合からカウント数を 0.6 倍する。同様に A.id のバケット 2 を基準にメモリサイズを計算し、上記の結果との合計が上限を超えているかチェックする。超えていた場合は、分割範囲として [10, 80] と [81, 120] の二つを設定する。

3. 検証試験

提案手法の有効性を検証するために、SQL クエリの実行メモリサイズを予測して分割範囲を計算し、予測した利用メモリサイズと実際の実行メモリサイズを比較した。ベンチマークとして TPC-H[3] のクエリ Q9 を利用し、データセットは Scale 50 を利用した。lineitem を DB1 に、それ以外を DB2 に配置し、クエリプランと処理順序及び処理場所を図 3 のように想定した。なお、MDBS における上限メモリサイズは 2GB とした。本環境において、MDBS 上で処理される各結合処理の分割範囲を計算した。その結果、分割基準属性は、各結合処理の結合条件に利用された属性が選択され、図 3 の結合処理 a において 3 分割、b において 2 分割、c において 2 分割となる分割範囲を計算することができた。一方結合処理 d は分割が発生しなかった。

本結果を利用して実際に分割実行した場合の利用メモリサイズと、提案手法の分割範囲計算時の予測メモリサイズを比較したグラフが図 4

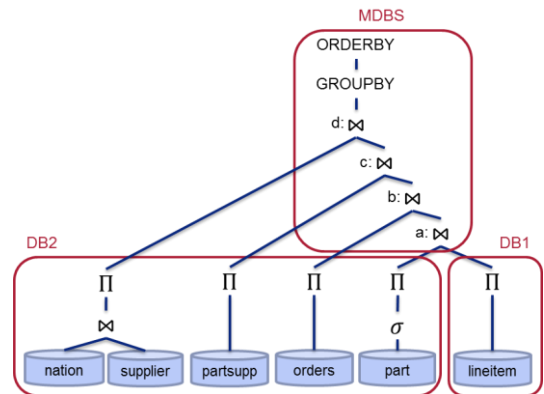


図 3 TPC-H Q9 のクエリプラン

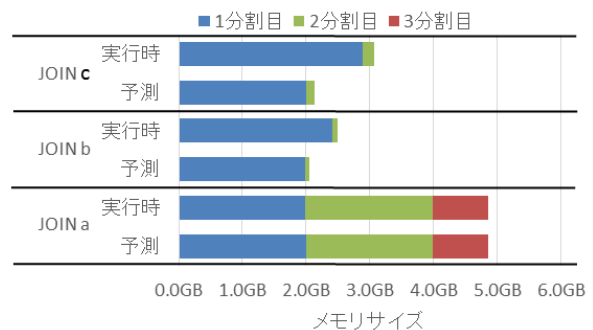


図 4 各結合処理の分割クエリ毎のメモリサイズ

である。提案手法は設定通り 2GB 以下で分割範囲を計算していることがわかる。しかし各処理の実行時のサイズと比較すると、b の 1 分割目で約 400MB、c の 1 分割目で約 900MB の差が発生している。これは結合処理における出力データのヒストグラム作成 [2] において、バケット内で結合されるドメイン数を少なく見積もった為である。そのため、提案手法を用いてクエリを分割しても、利用メモリサイズが実行時に上限を超える可能性がある。このことから、結合処理における出力データのヒストグラムの作成の改善や、誤差を考慮した分割範囲設定が必要と考えられる。

4. おわりに

本稿では、MDBS のメモリサイズを超えずに SQL クエリを実行するために、DBS へ投稿する SQL クエリを分割する範囲を、ヒストグラムを用いて計算する手法を提案した。引き続き分割範囲の精度向上を進めたい。

参考文献

[1] M. T. Özusu and P. Valduriez. Principles of Distributed Database Systems, Third Edition, Springer (2011).
 [2] N. Bruno and S. Chaudhuri, Exploiting Statistics on Query Expressions for Optimization, in Proceedings of SIGMOD, pp.263-274 (2002).
 [3] TPC-H, <http://www.tpc.org/tpch/>