様式Ｃ－１９

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成２４年　６月　８日現在

機関番号：８２６２６
研究種目：研究活動スタート支援
研究期間：2010～2011
課題番号：２２８０００８６
研究課題名（和文）　大規模科学技術データのための分析データベースシステムの研究開発

研究課題名（英文）　Development of an Analytical Database System for Large Scientific Data

研究代表者
油井　誠　（YUI MAKOTO）
独立行政法人産業技術総合研究所・情報技術研究部門・研究員
研究者番号：１０５８６７１２

研究成果の概要（和文）：科学技術データの急激な増加は、トランザクション処理中心の現在のデータベースシステムに大きな課題をもたらしている。本研究で、我々は巨大科学技術データのための分析データベースを開発した。評価実験により、提案手法が MapReduce に基づくシステム（Hive）と TPC-H SF=100 で比較して、最大 22.3 倍、平均 8.97 倍優れた性能を示すことを示した。

研究成果の概要（英文）：Rapid growth in the volumes of scientific data has led to unprecedented challenges in current data management systems specialized for transactional processing. In this research, we have developed an analytical database system for large scientific data. The results of experimental evaluation showed that our system is much faster (up to 22.3x and 8.97x on average) than a MapReduce-based system (Hive) on TPC-H SF=100.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
２０１０年度	1,260,000	378,000	1,638,000
２０１１年度	1,160,000	348,000	1,508,000
年度			
年度			
年度			
総　計	2,420,000	726,000	3,146,000

研究分野：データベース
科研費の分科・細目：メディア情報学・データベース
キーワード：データベース管理システム、e サイエンス、並列データベースシステム、分散処理システム

１．研究開始当初の背景

天文学、生命科学、地質といった科学技術分野で、国際的な情報共有が進んでいる。今後、学際的に巨大データを扱うデータ指向の科学において、蓄積された膨大な科学技術データの中から有益な知見を見出すための高度なデータ分析技術の重要性が増すのは必至である。しかし、従来のデータベースシステムは細かい粒度のトランザクション処理を想定して設計されているため、バッチ的に複雑なデータ分析処理を扱うのは不向きである。現状のオープンソースの関係データベースは扱えないような 100GB 以上～ペタバイトの範囲のデータを扱うことは困難である。科学者が大規模データから有益な知識を得るためには、テラ～ペタバイト級の科学技術データを扱うことが可能な分析データベースシステムの整備が急務であった。

２．研究の目的

本研究の目的は、テラバイト以上の科学技術データを効率的に扱うためのデータ分析基盤を実現することである。特に．科学技術データを適用対象とし、地質分野との分野融合プロジェクトへ最終的に適用することを目指している。

３．研究の方法

平成 22 年度より、研究代表者が研究開発中の汎用の分析データベースを科学技術データベース用途に移植した。本研究では、地理情報の大規模データ管理に、開発成果を最終的に適用することを目指している。そこで、ジオメトリ演算をサポートする関係データベース MonetDB を基礎として、MonetDB を無共有並列型の並列データベースに拡張することで、地理情報データのデータ量の増加に対してスケール可能なシステム基盤を開発した。

本課題では、最終的な目標である分野融合プロジェクトへの適用に向けて、開発したシステムの実装に対して 2 つの予備的な評価を行った。

（１）33 台程度の小規模な計算機クラスタ環境において、MapReduce に基づく競合システム Hadoop/Hive と性能比較を行った。

（２）分野融合プロジェクトへの適用については、産業技術総合研究所の地球観測グリッド研究グループ が参加する月周回衛星「かぐや」(http://www.kaguya.jaxa.jp/) のスペクトルデータの解析に本研究の成果を適用することを目標に、予備的な評価を行った。

要約すると、（１）でデータ量に対するスケーラビリティなどデータベースとしての基本的な性能の検証を行い、（２）で月周回衛星の実用化のために実データを用いた簡易評価を行う。

４．研究成果

大規模データを扱う上ではデータの事前分割に基づく並列処理が鍵となる。そこで本研究では、適応的なデータ交換を低減する事前データ分割手法（φハッシュ分割）と主記憶データベースの組合せによる主記憶を有効活用した MapReduce 処理手法を開発した。

従来のテーブル分割手法は、テーブルの一つに属性値に基づいてデータの分割を行う。これでは、複数の属性値を利用した複数の結合演算において、データの再配置が必要となるという問題があった。

具体的な問題点を、3 つのテーブル R1、R2、R3 の結合演算「R1.A = R2.A ∧ R2.B = R3.B」を例に述べる。これは、テーブル R1 と R2 が属性 A で結合され、テーブル R2 と R3 が属性 B で結合されるような問合せである。最初の結合演算「R1.A = R2.A」では、R2 が属性値 A に基づいてデータ分割されている必要がある。一方で、「R2.B = R3.B」では、R2 が属性値 B に基づいてデータ分割されている必要がある。つまり、ここで上記二つのデータ分割要求には矛盾が生じている。

提案手法では、従来手法とは異なり複数の属性値に基づいてテーブルのデータ分割を行う手法を採用した。その上で、どの属性値に基づいてその行が分割されたかの情報を行ごとに差し込む。

図 1 に例を示す。例えば、Key1 で分割されている行を赤で示し、Key2 で分割されている行を緑とする。

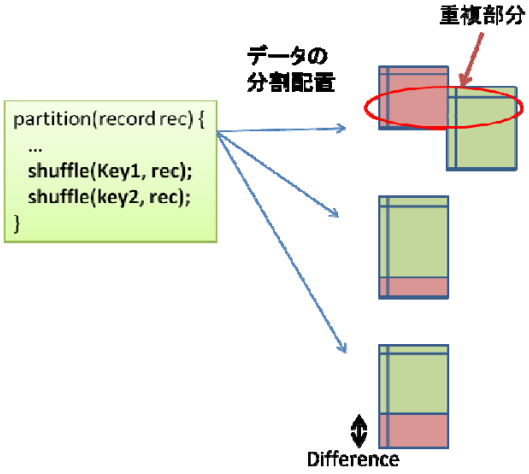

図 1. 複数のキーに基づいたデータ分割

複数のキーに基づいて分割することで、例えば Key1 で分割されていることを要求される問合せに対しては、赤い部分を、Key2 で分割されていることを要求されている問合せに対しては緑の部分を選択するような選択演算を問合せに追加することで、問合せが要求するデータ分割に対応することができる。

複数の属性値に基づいた結合演算に、この手法は応用ができる。「R1.A = R2.A ∧ R2.B = R3.B」を例にとれば、最初の結合演算「R1.A = R2.A」では R1、R2 をそれぞれ属性値 A に基づいて分割されたものとする選択演算を加える。その上で、次の結合演算「R2.B = R3.B」で、R2、R3 がそれぞれ属性値 B に基づいて分割されたものとする選択演算を加えればよい。

このφハッシュ分割の導入の狙いは、データの移動をなくしてディスク I/O やネットワークの負荷といった shuffle 操作のオーバヘッドを下げることだけでなく、データの移動にともなうメモリからのデータ入出力を避け、無共有型並列データベースの各データベース

ノードで問合せを極力インメモリ（memory mapped データに対する map 処理と単一の reducer による処理）で処理することにある。

これまでに得られた定量的・具体的な成果は、参考文献 1 で TPC-H と呼ばれるデータウェアハウスの業界標準ベンチマークのスケールファクタ 100（100GB のデータセット）による評価である。無共有型の 32 ノード構成を用いて提案手法を実装（MonetDB/MR）したとき、データの再分割が必要となる競合手法（Hadoop/Hive）に対して 10 倍以上の性能を示すことがあることが検証できた。

データ移動のない特徴からメモリ上でのデータ処理が有効であるため、これを実現する改良を行い従来手法に比べて約 20 倍から 250 倍の高速化を達成した。表 1 に、TPC-H の 22 個の問い合わせについて、競合実装である Hive との比較結果を示す。

Query	MonetDB/MR (disk)	MonetDB/MR (memory)	Hive
Q1	159.2	5.9	1283.9
Q2	41.2	5.8	311.6
Q3	97.6	14.3	295.3
Q4	78.5	5.5	228.5
Q5	85.6	17.4	439.9
Q6	111.1	4.6	104.1
Q7	158.7	38.3	950.3
Q8	141.6	11.3	495.5
Q9	204.2	31.8	858.9
Q10	679.1	137.0	480.6
Q11	59.4	5.0	248.3
Q12	118.8	12.0	165.5
Q13	197.7	42.5	330.5
Q14	79.5	6.0	127.5
Q15	127.2	4.8	216.8
Q16	36.6	17.2	338.9
Q17	92.8	6.8	336.0
Q18	188.8	67.3	471.8
Q19	82.2	10.4	249.0
Q20	135.9	2.9	519.5
Q21	133.1	28.6	916.3
Q22	92.0	13.8	357.4

表 1. TPC-H SF=100 による競合実装との性能比較（単位は秒）

本成果は雑誌論文①に発表した。鍵となったのは、ジョイン演算処理時の適応的なデータ交換を低減して、データ並列の問合せ処理を可能としたことであり、これを実現したデータ分割配置手法を新規に開発して特許出願を行った。既に、この特許出願技術をもとに企業との共同研究を実施しており、今後の展望としては、当該技術の普及のための整備、開発を行っていく予定である。

また、提案システムを地質分野との分野融合プロジェクトに適応するにあたって、所属機関のおける地質情報処理の専門家と連携して、月周回衛星「かぐや」の衛星データ活用における典型的なデータベース問合せの抽出やデータベースの利用パターンの同定を行った。かぐやの生データ 4.5TB であるが、予備実験では、この一部を評価データセットとして利用した。これまでに、既存システムにおける問題点（データのローディング時間がネックとなること）を抽出したとともに、提案システムにデータの一部を投入したときに、提案システムで問合せ処理が可能なことまでを確認した。今後、4.5TB の全量を提案システムに投入しての評価を進めていく予定である。

５．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計１件）

① 油井誠, 小島功、タプル再分散不要の並列データベース構成法、情報処理学会論文誌：データベース, Vol. 4, No. 4, (TOD52), pp. 11-33, 情報処理学会, 2011 年 12 月.査読有
http://id.nii.ac.jp/1001/00079662/

〔学会発表〕（計１件）

① Makoto Yui and Isao Kojima、SQLET: A Database Programming Language and Execution Environment for Parallel SQL Processing running on Plain RDBMSs、第 4 回データ工学と情報マネジメントに関するフォーラム（DEIM2012）論文集D2-5, 2012 年 3 月.査読無、シーサイドホテル舞子ビラ神戸（神戸市）
http://db-event.jpn.org/deim2012

〔産業財産権〕
○出願状況（計１件）

名称：表データのデータ処理方法、データ処理システムおよびそのコンピュータプログラム
発明者：油井　誠
権利者：独立行政法人産業技術総合研究所
種類：特許
番号：特願 2010-232069
出願年月日：平成 22 年 10 月 15 日
国内外の別：国内

６．研究組織

(1)研究代表者

油井　誠（MAKOTO YUI）
独立行政法人産業技術総合研究所・情報技術研究部門・研究員
研究者番号：10586712