

論文 / 著書情報
Article / Book Information

題目(和文)	高い鮮度で一貫性保証と更新集約適用を行う HTAPシステム高性能化に関する研究
Title(English)	A Study on Performance Enhancement of Highly Data-Fresh HTAP Systems Guaranteeing the Consistency and Aggregating Updates
著者(和文)	塩井隆円
Author(English)	Takamitsu Shioi
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12816号, 授与年月日:2024年6月30日, 学位の種別:課程博士, 審査員:宮崎 純,吉瀬 謙二,小林 隆志,渡部 卓雄,金子 晴彦
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12816号, Conferred date:2024/6/30, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

論文要旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	情報工学 情報工学	系 コース	申請学位 (専攻分野)： 博士 (工学) Academic Degree Requested Doctor of
学生氏名： Student's Name	塩井 隆円		指導教員 (主)： 宮崎 純 Academic Supervisor(main)
			指導教員 (副)： Academic Supervisor(sub)

要旨 (和文 2000 字程度)

Thesis Summary (approx.2000 Japanese Characters)

オンラインに生成されるデータの急速かつ継続的な増加により、データベースシステム (DBMS) に最新データを高速に更新するだけでなく、蓄積した大量データを効率良く分析することが伝統的に求められている。従来、データベースシステムにおけるデータを更新する処理 online transactional processing (OLTP) と、データを分析する処理 online analytical processing (OLAP) 其々のワークロード特徴に合わせて OLTP システムから OLAP 専用システムへ分離して最適化・開発されてきた。OLTP システムのデータを OLAP システムへ反映するために定期的にバッチ更新処理が行われてきたが、リアルタイムに更新された最新データを分析に利用できないことが問題とされていた。そのようなデータフレッシュネス問題を解決するために、近年、一つのシステムで OLTP と OLAP を効率良く実行するための Hybrid transactional/analytical processing (HTAP) システムが大きな注目を集め広く利用されている。

HTAP システムの多くは最新データを OLAP 側で高速に読むためにデータの一貫性を犠牲にしている。従来、OLTP システムではデータの一貫性を保証する仕組み Concurrency Control (CC) の一つである two phase locking (2PL) などが利用されていたが、OLAP では大量のデータを分析する際にテーブル全体をスキャンするため 2PL など OLTP 側の書き込みがブロックされる。一般的に OLTP では発行されるトランザクションを短い実行時間で素早く終了・応答させるため、OLAP 側の一貫性保証によって大量の OLTP トランザクションが実行できず従来の OLTP 性能が大きく劣化することになる。よって、OLAP 側の一貫性保証をしつつ従来の OLTP 性能を維持できる仕組みが今までなかったため、HTAP システムはデータの一貫性を犠牲にして高いデータフレッシュネスが求められてきた。また、HTAP システムでは OLAP 高速化のために OLTP 性能が犠牲になりやすい。DBMS では一般的に表形式でデータを扱い、OLTP システムでは行単位、OLAP システムでは列単位とデータを連続してアクセスする仕組みが異なる。OLTP システムは正規化されたテーブルを、行単位に連続してデータを格納する行指向データフォーマットによって必要な行データのみを取り出して読み書きできる。一方で、データ分析は OLTP の更新処理とは異なり、表形式のデータの中から必要な列のみを使用する。そのため、OLAP システムでは不要な列データを除いた分析用テーブルを新たに作成し易い仕組みや、列単位で連続してデータを格納する列指向データフォーマットが採用されてきた。それゆえ、行指向データフォーマットのままでは OLAP のシステムリソースを消費しやすく、HTAP システムでは OLAP 性能劣化の影響によって、従来の OLTP 性能も大幅に劣化することを避けるために列指向データフォーマットも利用せざるを得なかった。しかしながら、列指向データのままでは OLTP の行単位の更新が遅くなるトレードオフがあるため、従来まで HTAP システムでは OLAP 高速化のために OLTP 性能が犠牲になってきた。

本論文では、HTAP システム全体のデータ一貫性を保証しつつ OLTP・OLAP 性能を犠牲にしない新たな CC 手法の Read-safe snapshot (RSS) とその CC におけるバージョン識別子に基づいて OLTP の更新データを集約して列指向データに反映する仕組みの SCMAT を提案する。

従来の CC 手法では、HTAP 全体の性能が大幅に劣化し一貫性保証可能なリアルタイムデータ分析基盤の実現は難しい。伝統的に OLTP システムの CC 手法は研究されてきたが、HTAP のようにテーブルデータ全体に影響する OLAP が参加した場合に従来の CC 手法を用いると OLTP や OLAP で大量のトランザクションが abort/wait され、HTAP 全体の性能が大幅に劣化する。RSS は、HTAP 用ワークロード CH-BenCHmark において、従来手法 SSI-SafeSnapshot よりも約 45% OLTP 性能を向上させつつ OLAP 性能の劣化がなかった。

RSS は同じデータの複数バージョンを利用する Multi-version CC において、不整合を起こさない最新データを読むための新たな理論的枠組みである。HTAP システムにおいて直近で書かれた最新のデータは、並行して実行される大量トランザクションの影響によって OLAP 側で読み出すと anomaly が起こる可能性が高い。本論文ではそのような anomaly 構造を read-only anomaly として定義し、CH-BenCHmark では約 10% の高い確率で起こることを確認した。RSS を使用した場合、read-only anomaly 群が起こらないことを理論的に証明し、マルチノード環境の HTAP システムで実験した場合も read-only anomaly cycle が起こらず、代表的な HTAP システム CochroachDB を利用するよりも全体の OLTP/OLAP 性能が大きく向上した。

従来の HTAP システムでは OLAP 側に最新の更新データを含めて読めるようにするために、リアルタイムに OLTP の細かい更新差分を列指向データへ反映する処理を待つ必要があり、データの反映処理を待つ時間が OLTP 性能を劣化させる。SCMAT は更新差分の行データのバージョンを識別する情報と列指向データのブロック識別子をメモリ上のインデックスに管理させ、更新頻度の高い列指向データへ優先的に反映する仕組みである。OLAP システムの MariaDB columnstore 上に実装し、TPC-C ベンチマークテストによる実験で OLTP 性能が約 95% 向上した。

本論文の理論と手法は、高性能かつ一貫性保証可能なリアルタイムビッグデータ分析基盤のシステムで実際に適用されている。

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note : Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1 copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).

(博士課程)
Doctoral Program

論文要旨

THESIS SUMMARY

系・コース： Computer science 系
Department of Graduate major in コース
学生氏名： Takamitsu Shioi
Student's Name

申請学位 (専攻分野)： 博士 (Engineering)
Academic Degree Requested Doctor of
指導教員 (主)： Jun Miyazaki
Academic Supervisor(main)
指導教員 (副)：
Academic Supervisor(sub)

要旨 (英文 300 語程度)

Thesis Summary (approx.300 English Words)

This dissertation describes on the theory and methodology of a high-performance serializable Hybrid transactional/analytical processing (HTAP) system for consistent real-time data analyses. HTAP systems developed to efficiently execute two types of processing in the database management system (DBMS), online transactional processing (OLTP) mainly for write and online analytical processing (OLAP) mainly for read of data, have attracted much attention.

Concurrency control (CC) that ensures consistency/serializability of updated data is an essential element of database management system (DBMS).

Even if only the consistency of the transaction history on the OLTP side is guaranteed, the consistency of HTAP systems might be broken due to the participation of read-only transactions on the OLAP side. Systems need to allow many users to read/write data concurrently, so inconsistent results and anomalies will be obtained if consistency is not guaranteed.

Moreover, due to the property of connecting OLTP and OLAP components, HTAP systems need to deal with the problem of different access methods.

OLAP tends to access data represented in tabular form per column but OLTP handles data per row.

The conversion from processed data in OLTP to columnar data is the one of major issues in HTAP systems for accelerating OLAP.

This dissertation proposes read-safe snapshot (RSS) that ensures consistency and high performance, and a mechanism named SCMAT that aggregates the conversion from processed data in OLTP to columnar data based on a version identifier on CC.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).