

論文の内容の要旨

論文題目 異種データベースシステムの連携技術に関する研究

氏 名 池田 哲夫

データベース技術の発展・普及に伴い、既存の複数のデータベースシステムを連携して利用したいという要求が生じている。

ここで、既存の、互いに異種性を有する（即ち、データモデル、データ項目の値の表現形式などが同一とは限らない）データベースシステムを連携して利用する形態の代表的なものに、①基幹業務の効率化などを狙いとして基幹系データベースシステムの仮想的な連携を実現する連邦データベースシステムと、②データ入力稼働削減、組織内を流れる情報の速度の向上などを狙いとして基幹系データベースシステム間で直接的にデータを交換することにより連携を実現する企業アプリケーション統合と、③的確な意志決定を可能にすることを狙いとして基幹系データベースシステムからデータを抽出して構築される意志決定支援用のデータベースシステムであるデータウェアハウスとの3つがある（図1）。本研究では、各形態において重要と認識されている以下の問題点を研究課題として設定している。

- ・ 連邦データベースシステムにおけるスキーマの構築方式
- ・ 企業アプリケーション統合におけるデータ交換システムの構築方式
- ・ データウェアハウスにおけるオンラインデータローダの高速化方式

第1の研究課題（連邦データベースシステムにおけるスキーマ構築方式）に関して、研究内容を説明する。

スキーマとは、連邦データベースシステムに関する定義情報を言う。具体的には、連邦データベースシステム中のデータの構造、連邦データベースシステム中の実体・データ項目と基幹系データベースシステム中の実体・データ項目との対応関係、連邦データベースシステム中の各種の整合性制約、などからなる。

従来の研究においては、スキーマの構築稼働が大きいという問題があった。

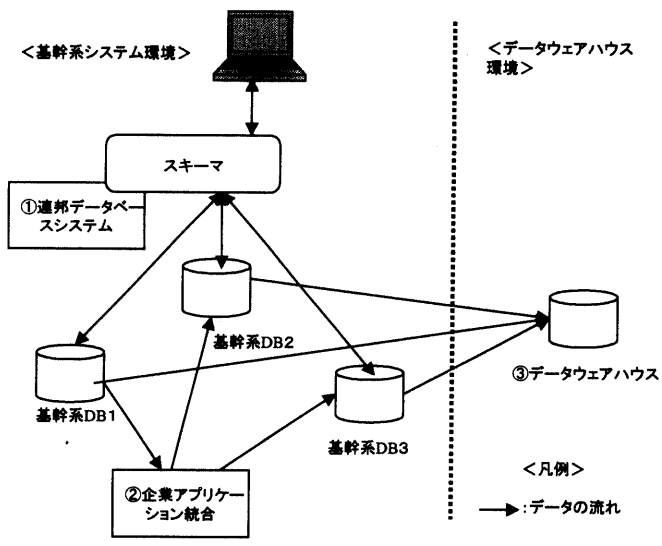


図1. 異種データベースシステムの連携の形態

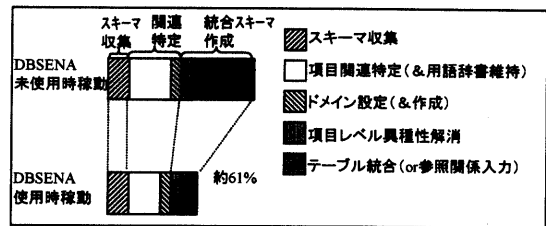


図2. スキーマ構築稼働の比較

この問題を解決するため、①実体レベルの作業を簡易化することを狙いに、関係データベースにおける質問の簡易化のために導入された概念である普遍関係を採用することと、②データ項目間の関連特定作業を簡易化することを狙いに、用語辞書を用いて類似データ項目の分類を支援する機能を備えることを主な特徴とするスキーマ構築方式を提案している。

提案方式を実現する連邦データベース検索システム DBSENA(DataBase SEmantic NAVigation system)を作成し、さらに DBSENA を用いて複数データベースシステムを連携するシステムを試作し、提案方式の有効性を評価した。

提案方式の特徴を有しないスキーマ構築方式（実体レベルの構築作業を含み、類似データ項目の分類支援機能を有しない構築方式）を用いて構築した場合と構築稼働の比較を行った（図2）。比較の結果、構築稼働が約4割削減されたことが判明し、提案方式は構築簡易化に有効であることが確認された。

提案方式の適用領域に関しては、以下に記す広い適用領域を有することを示した。

企業が情報資源管理（情報を企業経営資源の一つとみなして、その完全性を維持し、有効利用を図ること。情報資源管理の方法論の一部に用語の統制やデータ項目名の標準化などが含まれる）を実施しており、スキーマのデータ項目名の集合を各データ項目の意味が容易に理解可能なように構築できる場合の、総データ項目数が約200程度までのスキーマの構築。

（スキーマのデータ項目名の集合を各データ項目の意味が容易に理解可能なように構築できるか否かに関わらず）、総データ項目数が数十程度までのスキーマの構築。

第2の研究課題（企業アプリケーション統合におけるデータ交換システムの構築方式）に関して、研究内容を説明する。

データ交換システムとは、基幹系のデータベースシステム間での直接的なデータ交換を実現するシステムを言う。従来より存在する、①業務として定常的にデータ交換を行なうシステム、②情報システム更改時用

のデータ移行システム、③データ複製システム、などを包括する概念である。

データ交換システムの構築において、交換先データベースシステムの自律性（レガシーシステムであり既存機能の改造が困難である、など）に起因して、交換先の既存データ入力インタフェースを改造無しに利用する制約などの、交換先への影響を極小化する制約を課された環境が考えられる。従来の研究においては、この環境を含む様々な環境向けのデータ交換システムを、少ない稼働で構築可能とする構築方式の研究は無かった。

この問題を解決するため、①様々な環境向けのデータ交換に必要な機能の明確化を狙いに、データ交換処理モデルを提案し、②構築簡易化を狙いに、データ交換処理モデルに沿った、データ交換システムの開発・実行環境を提案している。

データ交換処理モデルは、データと機能との形式的なモデルの考案と、それに基づく操作機能の導出とを特徴とする。

データ交換システム開発・実行環境は、生成方式を適用すること、即ち、データ交換の仕様を簡易に記述できる高水準仕様記述言語を備えることと、基本的な操作機能およびデータ項目値変換関数をソフトウェア部品として備えることとを特徴とする。

データ交換システム開発・実行環境の実装（DB-STREAM と呼ぶ）を行い、さらにDB-STREAM を用いて複数データベースシステム間のデータ交換システムを試作し、データ交換処理モデルおよびデータ交換システム開発・実行環境の有効性を評価した。

交換先データベースシステムの既存データ入力インタフェースを改造無しに利用するデータ交換システムが構築でき、データ交換処理モデルは交換先への影響の極小化に有効であることが確認された。

データ交換システム構築稼働に関しては、個別プログラミングによって構築した場合と構築稼働の比較を行った（図3）。比較の結果、構築稼働が約6割削減されたことが判明し、データ交換システム開発・実行環境は構築簡易化に有効であることが確認された。

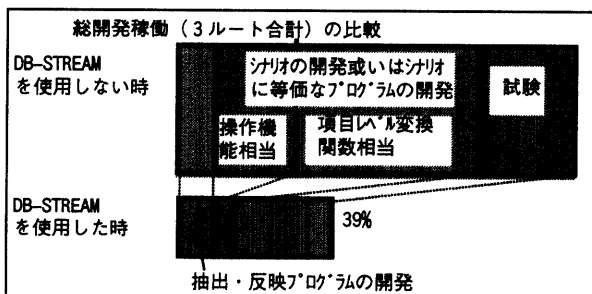


図3. データ交換システム開発稼働の比較

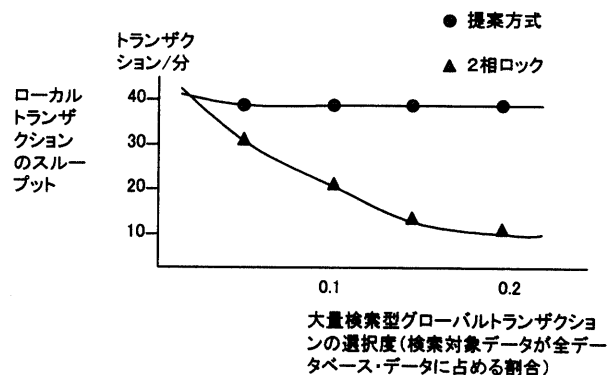


図4. 同期制御方式の性能の比較

提案した開発・実行環境の適用領域に関しては、①既存データ入力インタフェースを改造無しに利用するなどの、交換先への影響を極小化する制約を課された環境でのデータ交換システムの構築が適用領域となることと、②前記制約を課されない環境でのデータ交換システムの構築においても、適用対象のシステムで必

要な機能の多寡に依存して、提案した開発・実行環境が最適となりうるケースが少なからぬ割合で存在することとを示した。

第3の研究課題（データウェアハウスにおけるオンラインデータローダの高速化方式）に関して、研究内容を説明する。

データローダとは、基幹系データベースシステムからデータを抽出してデータウェアハウスに該データをロードするシステムを言う。基幹系データベースシステムの大量のデータを検索することを特徴とする。データローダは、少量のデータの検索・更新を行なう通常のトランザクションとは非同期にロードを行なうオフラインデータローダと、トランザクション技術を用いて、通常のトランザクションと同期してロードを行なうオンラインデータローダとに大別される。

従来の研究においては、通常のトランザクションと、オンラインデータローダ用の大量の検索を行うトランザクション（大量検索型グローバルトランザクションと呼ぶ）が共存する環境における、高性能なトランザクション同期制御方式が確立していなかった。特に、ミッションクリティカルなものが多数存在する通常のトランザクションの高性能な実行を可能とする同期制御方式が確立していなかった。

この問題を解決するため、トランザクションが検索専用の場合には、同期制御に起因する副作用が小さい（アポートは一切発生せず、他のトランザクションとの間で待ちが発生するのは READ の実行時に、未完了のトランザクションによって作成された版を選択しそのトランザクションの完了を待つ場合のみ）という特長を持つ方式として多版時刻印方式があることに着目し、該方式と同期制御方式の主流である2相ロック方式とをベースに新たな方式を提案している。提案した方式は以下の特徴を有する。

- ・ 通常のトランザクションの READ/WRITE の制御には2相ロック方式を修正適用する。
- ・ 大量検索型グローバルトランザクションの READ の制御には多版時刻印方式を修正適用する。
- ・ 大量検索型グローバルトランザクションの WRITESSET（トランザクションでの更新対象テーブルの和）に関して、大量検索型グローバルトランザクションを他のトランザクションと直列に実行させる。

提案方式の性能に関して、従来方式（2相ロック方式）の性能との比較をシミュレーションによって行なった（図4）。比較の結果、通常のトランザクションの性能（スループット）は提案方式の方が大幅に（最大約3倍）良いことが判明し、提案方式は、通常のトランザクションと大量検索型グローバルトランザクションとが共存する環境における優れた同期制御方式であることが確認された。

提案方式の適用領域に関しては、遅延の少ない意志決定支援を可能にすることを狙いに、オフラインデータローダではなくオンラインデータローダが構築されるケースが少なからぬ割合で存在すること、従って提案方式が一定の適用領域を有することを示した。

以上を要するに、本論文では、異種データベースシステムを連携して利用可能とする技術を確立するために、種々の重要な問題の解決方式の提案を行い、稼動データあるいは性能シミュレーションデータの定量的な比較により、それらの提案方式の有効性を確認した。さらに、それらの提案方式が広い適用領域を有することを示した。