

論文内容の要旨

博士論文題目

Making XML Database Systems Scalable to Computer Resources and Data Volumes
(計算リソースとデータ量にスケーラブルな XML データベースの実現)

情報科学研究科情報生命科学専攻データベース学講座 油井誠

(論文内容の要旨)

XMLの利用の拡大により多量のXMLデータを効率的に管理することが求められている。データベースの利用者がデータベースに求める要件のうち、最も重要となる項目の一つが、データベースの計算リソースやデータ量に対して得られる性能のスケーラビリティである。本研究では、(a)データ量に対して得られる性能のスケーラビリティ、(b)無共有型計算機クラスタを利用したスケーラビリティ、(c)共有メモリ型並列計算機を利用したスケーラビリティの三つの観点で、それぞれXMLデータベースのスケーラビリティ向上のための手法を提案する。(a)のデータ量に対して得られる性能のスケーラビリティでは、表形式で内部表現されたXML文書をブロック化して二次記憶に文書順に配置し、問合せ実行中に必要なブロックを効率的に主記憶に読み込む機能を特徴とするXML問合せ処理手法を開発した。実験結果から提案手法がデータサイズに対してほぼ線形の性能が得られることを示した。(b)の無共有型計算機クラスタを利用したスケーラビリティでは、リアルタイム性が要求されるXML問合せに対して、計算機クラスタを利用した分割統治を行う場合を議論する。これまでの分散XMLデータ処理で常であった値渡しによるデータ交換には中間計算結果のシーケンスの一部の要素のみが利用される場合に不要な通信や計算が発生する問題や、オペレータ間の並列性が阻害されるという問題がある。この問題への解決策として、遅延評価を活かした参照渡しに基づく分散XML問合せ処理手法を開発し、実験結果により分散XMLデータベースシステムが参照渡しによるデータ交換を考慮に入れることの重要性を実証した。(c)の共有メモリ型並列計算機を利用したスケーラビリティでは、メニーコア計算機上でデータベース管理システムを利用する上での問題点として、バッファ管理モジュールの並行アクセス耐性の低さを取り上げ、ノンブロッキングのバッファ管理手法を提案する。64プロセッサを備えるSun UltraSPARC T2を利用した実験により、既存のバッファ管理手法が16プロセッサ以上のスケールしないのに対して、提案するバッファ管理手法が64プロセッサまで、プロセッサ数に応じた線形の性能が得られることを示した。最後に三つの提案手法を組合せてXMLネイティブデータベースシステム上に実現する上で得られた知見を述べる。