

**科学研究費助成事業 研究成果報告書**

平成 28 年 6 月 6 日現在

機関番号：13901

研究種目：若手研究(B)

研究期間：2012～2015

課題番号：24700087

研究課題名(和文) マルチコア環境における複数演算の並列評価による低遅延データストリーム処理

研究課題名(英文) Low-latency processing for data streams by parallelly evaluating multiple operators in multi-core environments

研究代表者

渡辺 陽介 (Watanabe, Yousuke)

名古屋大学・未来社会創造機構・特任准教授

研究者番号：80532944

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：センサーデータ等のように、時間の経過に伴って生成されるストリームデータが増加し、それらに対するデータ処理技術に注目が集まっている。ストリームデータ処理ではリアルタイム性を求められることが多く、処理遅延の削減は重要課題の一つである。本研究では処理遅延を減らすために、複数のCPUコアを用いて同一データに対する複数処理を並列評価する処理方式の実現を目的としている。並列評価する演算について、直積演算と選択演算だけでなく、M入力の結合処理を行えるようにした。また、従来の問合せ最適化方式が機能せず、提案方式が有効に働く状況の一つとして、実行時の情報源の動的選択を検討し、プロトタイプシステムを作成した。

研究成果の概要(英文)：Data streams which produce sequences of data according to time progress are increasing, thus data stream processing technologies have been focused. Since they often needs realtime processing, how to reduce latency caused by data processing is one of important research issues. The main purpose of this research is to achieve a parallel evaluation method for multiple operators by utilizing multi-core environments. The proposed method can treat not only simple queries such as product and selection operators, but also complex queries including M-way joins. And, the prototype system developed in this research supports dynamic information source selection, which cannot be covered by conventional query optimization methods.

研究分野：メディア情報学・データベース

キーワード：データストリーム マルチコア 並列処理 情報源選択

### 1. 研究開始当初の背景

近年、カメラ・センサーからの計測データや、Twitter からのつぶやきデータなど、時間の経過とともに大量に生み出され続けるストリーム型のデータが増加している。次々と新しいデータが得られるストリームに対して、連続的な問合せ処理を実現するストリーム処理エンジンの開発が行われてきている。ストリームデータのさらなる増加に伴い、より少ない遅延でより大量のデータを処理できる、効率的な問合せ処理技術が求められるようになってきた。

処理遅延は、データが演算を通過する際に発生してしまう不可避のものであるが、応答性の高さを求められるアプリケーションにおいては非常に重要な要素である。制御機器のような環境では、センサーデータを受け取ってからデータを加工して出力するまでの遅延時間に「ms 以内」などの制約がある場合がある。特に問合せが複雑化してくると、入力データが出力までに通らなければならない演算の数が増え、処理遅延の増加が問題となっている。

これまでの並列化手法は、スループットの向上、即ち CPU コアごとに異なるデータを同時処理させることで、システム全体として単位時間あたりに扱えるデータ量を増加させることに重点を置いてきた。一方で、入力されたデータをより少ない遅延で出力することに関してはあまり考慮してきていない。演算の評価順序を決定付ける処理プランは、逐次実行の時の処理プランをベースに構築しており、そのため入力データが通過しなければならない演算数は逐次実行の時と同じままであった。この状態で複数の CPU コアを使ったとしても処理遅延の削減効果に限界があった。

### 2. 研究の目的

本研究では、センサーデータ等を処理するためのシステムであるストリーム処理エンジンを対象に、マルチコア CPU 環境を活かして処理遅延時間を削減するための複数演算の並列評価手法を提案することが目的である。従来、個々のデータに対して順番に評価していくはずだった演算を別々の CPU コアに割り当て、同じデータに対する演算評価を同時並列的に行うことで処理遅延の短縮を狙った。

### 3. 研究の方法

本研究では、ストリーム処理エンジンにおける複数演算の並列評価を行う仕組みを実現するために、理論的な検証とプロトタイプシステム実装に基づく実践的な検証の両面からアプローチし、問題解決を図った。

まず、並列実行可能な演算と逐次実行が必要な演算をシステムが区別できるようにするため、様々な演算同士の組合せの並列実行可

能性について理論的な検討を行った。リレーショナル代数演算のうち比較的簡単な選択演算や射影演算等を対象に問題を解き、次に高度なウィンドウを扱う結合や集約演算を含めた場合について検討を行った。

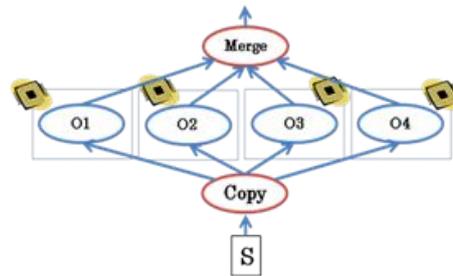
また、実際のストリーム処理におけるアプリケーション側の要求を分析し、処理遅延の少ない並列処理が必要となるクエリについて調査した。具体的にはストレージシステムの消費電力の計測ストリームや、移動物体の位置に応じた情報源の動的選択処理のシチュエーションを扱った。

最後に、より現実的な実験評価を可能とするため、実際のストリーム処理エンジンに提案手法を組み込んだものを開発した。

### 4. 研究成果

#### 基本戦略：

問合せ要求に含まれる演算のうち、依存関係のないものは並列評価できる点に着目する。下図の例では、演算 o1 から o4 に依存関係がない場合、各演算に同時に同じタプルをコピーし、最後に結果をマージするような処理プランが生成できる。演算を並列評価することにより、処理プランの木の高さ自体を減らすことができるため、処理遅延の削減が期待できる。これまでの並列処理方式では、CPU コアが並列に動くときは別々のデータを対象に処理を行っているが、提案手法では同じデータを複数の CPU コアに渡して別々の演算を処理し、最後にそれぞれの CPU コアの結果をまた 1 つのデータに戻す、という動作をする点が異なる。



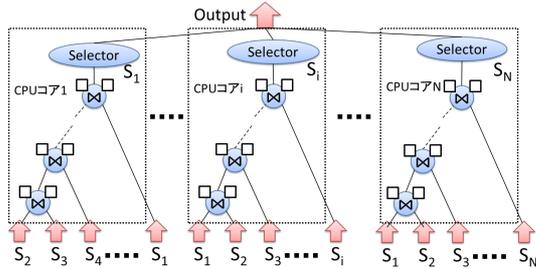
選択演算、射影演算、関数評価の単純な演算については、この方式で並列処理が実現できることを確認した。

#### M 入力の結合処理の処理遅延削減：

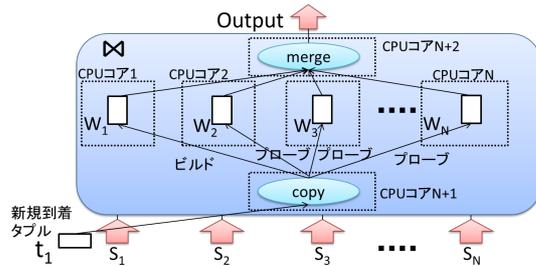
M 入力のストリームデータに対する結合処理に着目し、マルチコア CPU を活かした並列な結合処理方式を 2 つ提案した。

(1) 並列結合ツリー方式は、2 入力の結合演算をツリー状に組合せて M 個のストリームを結合する従来方式を拡張したものである。一般に、各ストリームの入力レートと各結合条件の選択率が完全に既知で、しかもそれらが時間変化しないという前提であれば、最適な結合順序を表すツリーを一意に求めることが可能である。しかし、データストリームでは

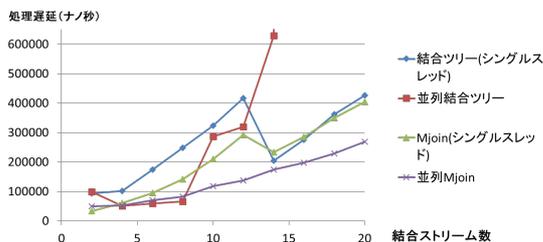
その仮定は成り立たないことが多い。そこで、結合順序の違いによって処理遅延がストリームごとに变化する性質に着目し、結合順序を変えた複数のツリーをマルチコアで複数同時並列的に評価することで、より少ない遅延で出てきた出力結果を毎回優先的に採用するという方式をとる(下図)。



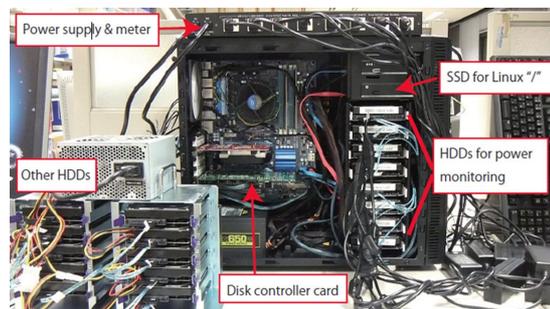
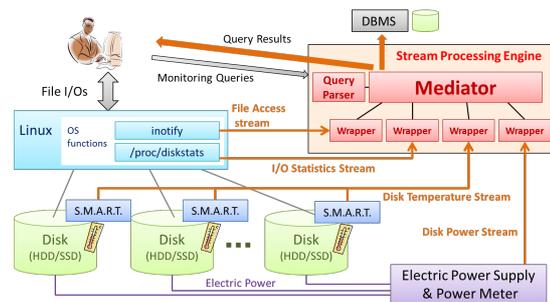
(2) 並列 MJoin 方式は、既存の M 入力の結合アルゴリズム MJoin をマルチコア向けに改良したものである。MJoin では、M 個の入力ストリームに対応するウィンドウ領域とハッシュ表を保持している。あるストリームから新規タプルが到着した場合、まず新規タプルをそのストリームのウィンドウとハッシュ表へ反映させるビルドの処理を行い、その後で他のストリームの各ハッシュ表とマッチングさせるためのプロブ処理を行う。1 つのハッシュ表のプロブ処理が終わったら、それを中間結果として次のハッシュ表のプロブ処理を行うことを M-1 回繰り返し、最後まで残ったものを出力する。一般的にはプロブ処理を行うストリームの順番は、結合条件の選択率等に応じて中間結果がより少なくなるようなものが選ばれる。本研究の目的である、処理遅延の削減という目的で着目したのは、シーケンシャルに行われているプロブ処理である。この M-1 回のプロブ処理をマルチコア環境向けに並列化したものが並列 MJoin 方式である(下図)。



(3) 評価実験として、人工データのワークロードを用いて、これら 2 手法と従来手法を比較し、特に並列 MJoin 方式がより多くの処理遅延を削減できることを示した(下図)。



**ストレージシステムの消費電力ストリーム:** ストリーム処理技術の実応用として、ストレージシステムにおける消費電力測定環境を構築した(下図)。ストレージシステム内のハードディスクの個別の消費電力が測定できるように電力計を接続し、消費電力ストリームデータを連続的に収集・分析できるようにした。この測定環境の構築により、ストレージ省電力化に関する研究の実証実験に貢献するだけでなく、応用側からみた場合のストリーム処理の処理遅延削減への要求事項についてもより明らかになった。また、提案手法の利点を主張するために、従来の間合せ最適化技術が使用できない状況を集中的に扱うべきであることがわかってきた。



**移動物体の位置に応じた情報源動的選択:** 従来からの間合せ最適化方式が機能せず、提案方式による並列化が有効になる利用状況の一つとして、動的に選択された情報源の統合について検討した。ここでいう動的な情報源の選択とは、ストリームデータの現在の値に依存して問合せ中で使用される情報源が決定されることであり、例えば「位置情報に基づいて最寄りのセンサーからの入力に切り替える」等の処理を想定している(下図)。

情報源が事前に特定できない要求例:  
「追跡対象者の移動に合わせて、距離10未満にあるカメラからの映像データを収集し続けてほしい」

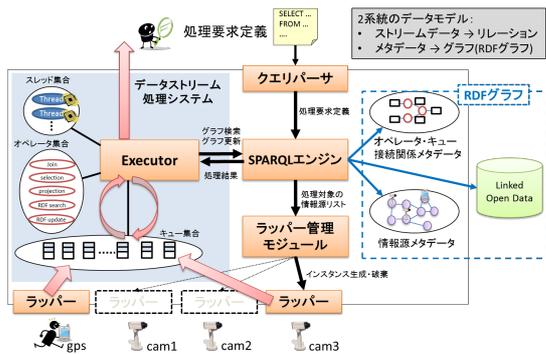


情報源が実行中に変更されることから、従来型の問合せ最適化方式でよく行われる、演算の選択率を事前に見積もり、最適な演算の実行順序を決定することが極めて困難である。それに対して、提案方式は複数のCPUコアによる演算の並列評価であるため、実行順序に依らない処理の効率化が可能である。この検討結果に基づき、本研究で実現するプロトタイプシステムに情報源の動的選択の機能を盛り込むため、必要な処理機構について考察・提案を行った。

### プロトタイプシステムの開発：

マルチコア環境を用いた並列処理手法の評価実験のために作成したプログラムをベースに、さらに情報源の動的選択機能を備えたストリーム処理エンジンの設計と実装を行った(下図)。

情報源の動的選択機能の設計を進めていく過程で、本機能の実現には実行中のクエリの内部構造を操作し、別のもの書き換える仕組みが必要であることがわかった。クエリの内部構造は、複数の演算やキューを連結したツリーデータまたはグラフデータとみなすことができる。そのため、グラフデータの管理機構をシステム内部に組み込むことで、クエリ内部構造の操作に高い汎用性と柔軟性を持たせられると判断した。そこでリンクドオープンデータの記述などに用いられる RDF (Resource Description Framework) を用いてクエリの内部構造を表現し、その書き換えを RDF データに対する更新操作として実現した。



下図中の処理要求は「ユーザの位置情報の変化に応じて最寄りのカメラからの映像に切り替える」ものである。3行目の SELECT-FROM 節は SQL ベースの構文で、カメラからの映像ストリームに対する処理を記述しているが、メタデータ変数として「?camera」と「?video」が用いられている。この変数の実際の値は、次の METADATA 節の中を繰り返し評価することによって取得している。METADATA 節内は、SPARQL をベースとした構文となっており、RDF データに対する検索条件を指定する。「ns:informationSource」は情報源を意味するクラス、「ns:locationX (Y)」は座標情報

に対応する。VALUES 節は SPARQL の構文からの拡張で、ストリームに対する SQL の処理結果を変数に関連付けることができる。ここでは GPS からの座標が変数「?tx」、「?ty」に関連付けられる。この記述例の処理では、GPS からのストリームデータが届くたびに、変数?tx, ?ty の値が更新され、RDF データの検索が発生する。情報源の検索結果を格納する?camera, ?video の値が前回の評価から変化した場合、接続する情報源が変化したことを意味する。このときシステムは条件に見合ったカメラへの接続を確立し、映像ストリームを受信する。

```

PREFIX rdf: <http://www.w3.org/...>
PREFIX ns: <http://...>

SELECT ?video
FROM ?camera [1 second]
METADATA {
  SELECT ?camera, ?video
  WHERE {
    ?camera rdf:type ns:informationSource .
    ?camera ns:locationX ?cx .
    ?camera ns:locationY ?cy .
    FILTER DISTANCE(?cx, ?cy, ?tx, ?ty) < 10 .
    ?camera ns:hasAttribute ?video .
    ?video rdf:type ns:VideoData .
  }
}
VALUES (?tx, ?ty){
  SELECT Gps.x, Gps.y
  FROM Gps [1 second]
}

```

Annotations in the image: 'メタデータ変数' points to PREFIX lines; 'ストリームデータの処理②' points to SELECT/FROM; 'メタデータの処理' points to METADATA; 'ストリームデータの処理①' points to VALUES.

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)

[学会発表](計 9 件)

渡辺陽介, 横田治夫, 「低遅延ストリーム処理のための結合演算並列実行方式」第 154 回データベースシステム研究発表会, Vol. 2012-DBS-154 No.10, pp. 1-8, 2012 年 8 月.

Yousuke Watanabe and Haruo Yokota, "Development of a Disk-Power Monitoring Environment Using a Stream Processing Engine", Proc. International Workshop on Streaming Media Delivery and Management Systems (SMDMS 2012), 査読有, pp. 351-356, Victoria, Canada, November, 2012.

Yousuke Watanabe, Hidetaka Kamigaito and Haruo Yokota, "Style-based Similarity Search for Office XML Documents", Proc. International Conference on Information Integration and Web-based Applications & Services (iiWAS 2012), 査読有, pp. 138-146, Bali, Indonesia, December, 2012.

小栗寛生, 引田諭之, 渡辺陽介, 横田治夫, 「ディスクの消費電力データストリームを用いたストレージの省電力化手法」情報処理学会第 75 回全国大会, 2013 年 3 月. 小栗寛生, 引田諭之, 渡辺陽介, 横田治

夫,「リアルタイム電力計測に基づくストリージの省電力化手法の実装と評価」DEIM Forum 2013, 2013年3月.

Yousuke Watanabe and Haruo Yokota, "Evaluation of Parallel Window-Join Algorithms for Reduction of Processing Delay", Proc. International Workshop on Streaming Media Delivery and Management Systems (SMDMS 2013), 査読有, pp. 635-640, Compiègne, France, October, 2013.

渡辺陽介, 横田治夫,「データストリームに対する連続的問合せの動的変更機能の提案」情報処理学会第76回全国大会, pp. 469-470, 2014年3月.

渡辺陽介,「データストリーム処理技術の研究動向について」,名古屋大学大学院情報科学研究科附属組込みシステム研究センター全体研究会(講演), 2014年6月.

Yousuke Watanabe and Haruo Yokota, "Dynamic Modification of Continuous Queries by Using RDF Metadata of Information Sources", Proc. International Workshop on Streaming Media Delivery and Management Systems (SMDMS 2015) in conjunction with 3PGCIC-2015, 査読有, pp. 754-759, Krakow, Poland, November, 2015.

## 6. 研究組織

### (1) 研究代表者

渡辺陽介 (WATANABE Yousuke)

名古屋大学・未来社会創造機構・特任准教授

研究者番号: 80532944