

次世代大規模分散・並列環境における高度メタデータ管理・解析システム技術の研究開発 (092103006)

Large-scale Distributed Metadata Management and Analysis System Technologies

研究代表者

谷村 勇輔 独立行政法人産業技術総合研究所

Yusuke Tanimura National Institute of Advanced Science and Technology

研究分担者

リンデン スティーブン[†] 的野 晃整[†]

Steven Lynden[†] Akiyoshi Matono[†]

[†]独立行政法人産業技術総合研究所

[†]National Institute of Advanced Science and Technology

研究期間 平成 21 年度～平成 23 年度 9 月末

概要

次世代分散並列基盤としての Google File System と MapReduce に基づいたアーキテクチャにおいて、メタデータ記述の枠組として注目を集める RDF データを大規模に管理・解析できる技術の研究開発を行った。RDF データの格納手法と並列処理、動的な問合せ・解析処理計画の最適化、効率的な中間結果の構築の 3 課題に取り組み、課題毎に大規模化に対応した新しい手法を提案するとともに各手法の連携方式の提案を行い、それらの有効性を示した。

Abstract

In this research, technologies to manage and analyze large-scale RDF datasets, on the architecture based Google File System and MapReduce, as a next-generation parallel and distributed computing environment, were studied. Three sub-themes were set in our approach: storage and parallel processing methods, dynamic query optimizations, and efficient intermediate data management for queries. Novel methods in each sub-themes and integration of them were proposed, and proved to be significant for large-scale RDF data processing, through our experiments.

1. まえがき

高度な知識処理ベースの構築には、様々なコンテンツやコンテキストを説明するためのメタデータが不可欠である。メタデータの表現方法として RDF (Resource Description Framework) は有望な技術の 1 つとして考えられており、W3C において標準が定められ、セマンティック Web で知られる知的処理や知的情報に広く用いられつつある。しかし、実用化に向けては情報量に比例して増加する膨大な RDF データの管理と高速処理基盤の構築が重要な課題となっている。特に、RDF の構造はラベル付き有向グラフであり、RDF データ処理には部分グラフ検索や経路探索等、コストの高い処理が求められる。また、RDF データの単純検索だけでなく、複雑な問合せ処理や集計処理、ルール処理、さらには RDF データが指し示すコンテンツも含めた大規模な解析処理が必要である。

一方、近年は Google File System や MapReduce といった新しい分散・並列処理基盤により、大規模データを扱う技術の進歩、普及が著しい。それらの技術を実装した Hadoop を中心としたオープンソースソフトウェアの開発も盛んである。これらを RDF データ処理に適用できれば、RDF データの実用化に大きく近づくことが可能である。

そこで、本研究ではそうした分散・並列処理基盤を RDF データ処理に適用するのに必要な技術を洗い出し、それらの開発を行った。

2. 研究内容及び成果

本研究の全体図を図 1 に示す。本研究では Pig/Hadoop により構築したデータ処理基盤を採用し、その上で大規模な RDF データを扱うのに必要な以下の 3 つの要素技術と

その統合技術を開発した。

- A) データ格納方式と並列データ処理の高速化手法
- B) 動的な問合せ・解析処理の最適化手法
- C) 効率的な中間結果の構築手法

これらはいずれも RDF データの特徴を利用した手法である。以降では手法毎にその概要と成果を述べる。

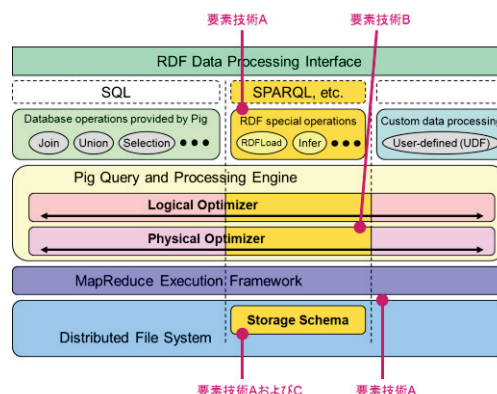


図 1 提案する大規模 RDF データ処理基盤

2.1 データ格納方式と並列データ処理の高速化手法

RDF データの特徴を利用した分散データ格納方式として、Vertical Partitioning をもとに Predicate 毎にファイルを作成し、各ファイルを MapReduce で扱いやすいバイナリ・データ格納方式 (MapFile) で管理する手法 (VP-MapFile) を提案した。そして、Pig から VP-MapFile を読み込むための RDFLoader モジュール、および Hadoop の分散ファイルシステム内に保存した既存の RDF フォー

マートを VP-MapFile に変換するツールを開発し、図 1 の基盤システムにおいて VP-MapFile を利用可能にした。

その上で、RDF データ処理の性能評価と高速化を行った。性能評価には RDF データのための SP²Bench ベンチマークを用い、SPARQL で記されたテストスクリプトを Pig Latin に翻訳して実験を行った。そして、既存の単一サーバによる RDF データベース (ARQ, メモリベースの Sesame, ディスク上のデータベースをバックエンドに用いた Sesame, Virtuoso) では 30 分以内に結果を得られなかった規模の RDF データ (2500 万トリプル) に対する問合せ処理 (q5a, q5b, q6) において、96CPU コアのクラスタを用いて、提案手法では数百秒で結果を得られることを示した。また、RDF データ処理において今後重要となるであろう Reasoning 処理に着目し、Transitive Closure 処理を Pig Latin を用いて実装し、結合演算のアルゴリズムや MapReduce のパラメータ・チューニングによる処理の高速化の指針を明らかにした。

2.2 動的な問合せ・解析処理の最適化手法

複数の RDF データリポジトリ (SPARQL Endpoint) がある場合に、各 Endpoint のメタ情報を使わずに、特定の Predicate に関する情報のみを利用して、各 Endpoint への問合せ計画を実行し、各結果の結合演算を受動的に最適化する手法を開発した。本手法は ADERIS ソフトウェアに実装し、<http://code.google.com/p/sparql-aderis/>にてソースコードを公開している。

また、ソーシャルネットワークの解析等、RDF データ処理では多段の結合演算が頻出する。これを MapReduce において高速実行するため、複数の結合演算を 1 度の MapReduce で行う手法が提案されているが、それを受動的最適化によりさらに高速化する手法を提案した。それは図 2 に示すように、最初の結合演算においてタスク間で演算の進捗や中間結果の情報を共有し、動的にタスク間の処理バランスや結合演算の順序変更を行うものである。この手法を実際に Hadoop の MapReduce に実装し、232CPU コアのクラスタを用いてその有効性を明らかにした。

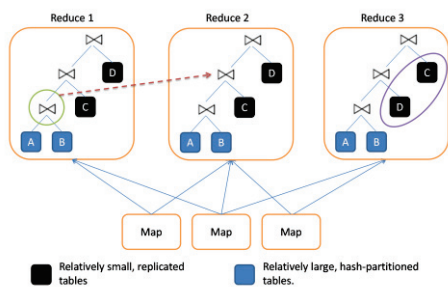


図 2 MapReduce における受動的な多段結合演算 (破線矢印：中間結果の転送、楕円枠：結合演算の順序変更)

2.3 効率的な中間結果の構築手法

RDF において頻出する結合選択率の低い結合演算のための結合アルゴリズムと、RDF 文書の意味のある纏まりに着目した格納法を開発した。前者では、B+木に対し Bloom フィルタを中間ノードに有するように拡張し、拡張 B+木を 2 つ突き合わせて、探索しながら非結合な枝を読み飛ばす手法を提案した。図 1 は提案手法の結合演算と従来手法であるソートマージ結合の結合処理時間を比較したものであり、提案手法では結合選択率が低い時に最大 420%ほどの高速化を達成した結果を示している。

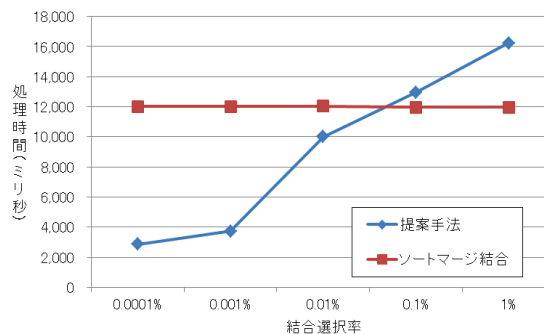


図 3 提案した結合演算アルゴリズムの性能

後者では RDF 問合せで頻出する高コストな結合演算のため、事前に結合した状態を保存する際に、RDF 文書のトリプルの順序を考慮し、ある程度の纏まりに結合して格納する手法を提案した。性能評価により、従来の 4 手法 (Triple Store, Vertical Partitioning, Same Subject, Flex Table) に比べて、提案手法が十分に高速であることを明らかにした。

3. むすび

本研究の成果としては、高度な知識処理ベースの構築に向けて大規模な RDF データを扱うための基盤を示し、その要素技術を開発したことである。Pig/Hadoop の枠組自体は汎用的なものであり、RDF データを効率的に扱うためには各レイヤにおいて様々な手法を開発する必要がある。そして、A や B の手法に関しては実際に Hadoop 上での実装・評価を行い、提案する基盤システムのスケラビリティの高さや各手法の連携による効果を明らかにした。また、C の手法に関しては今後、提案する基盤システムへの適用を開始するのに十分な基礎成果を得た。

本研究の成果により、従来の RDF データベースで扱えなかった規模の RDF データを扱えるようになり、今後 RDF データの一層の利用が進むとともに、RDF データを用いた情報の高度利用や知的処理を実現する知識処理ベースの構築が大きく前進することを期待している。

【誌上発表リスト】

- [1] Yusuke Tanimura, Akiyoshi Matono, Steven Lynden, Isao Kojima, “Extensions to the Pig Data Processing Platform for Scalable RDF Data Processing Using Hadoop,” Proceedings of 1st International Workshop on Data Engineering meets the Semantic Web (DESWeb) in conjunction with 26th IEEE International Conference on Data Engineering (ICDE), pp251-256 (March 6, 2010).
- [2] Steven Lynden, Yusuke Tanimura, Isao Kojima, Akiyoshi Matono, “Dynamic Data Redistribution for MapReduce Joins,” Proceedings of the Network Infrastructure Services as part of Cloud Computing (NetCloud) in conjunction with the CloudCom 2011, pp.717-723, (November 29, 2011).
- [3] 野見整、リンデン・スティーブン、谷村勇輔、小島功、 “BF マージ結合：Bloom フィルタを保持する B+木を用いた低結合選択率に適した結合手法”、情報処理学会論文誌：データベース (TOD)、Vol.4、No.2、pp114-125 (2011 年 7 月 1 日)。