

グリッド環境でのRDFデータベースの研究開発

セマンティックWebとグリッドを結びつけるデータベース技術



小島 功

こじま いさお

isao.kojima@aist.go.jp

グリッド研究センター
データグリッドチーム
研究チーム長
(つくばセンター)

グリッド基盤上でのデータベースの分散統合、特にWebDBやRDFなど異種のデータベースの統合を研究開発しています。地球観測データの分散統合基盤(GEO Grid)や、次世代のサービス基盤(AIST-SOA)におけるデータ管理機能などへ応用しながら、高度な分散知的基盤の実現を目指しています。

関連情報:

● プロジェクトホームページ
<http://www.dbgrid.org/>

共同研究者

木本正裕 (ビジネスサーチテクノロジー社)、Said Mirza Pahlevi、的野晃整 (産総研)

XMLを超えて

コンピュータに関わる多くの分野では、データをXML(eXtensible Markup Language)で記述することが一般となり、構造を持ったデータを自由に作成、交換できるようになりました。しかし、XMLで記述されるデータでは、例えば「植物は生物であるが、動物ではない」といった、データの性質(意味)をうまく表すことができません。このために、RDF(Resource Description Framework)という記述形式があり、セマンティックWebで知られる応用分野では不可欠な基盤となっています。

セマンティックWebとグリッド =セマンティック・グリッド

セマンティックWeb分野の発展に伴い、こうしたRDFのデータは増加する一方で、世界中で多様な形で生成・蓄積されています。さて、「リングは植物である」と前に述べた情報を組み合わせると「リングは動物ではない」が分かるように、RDFのデータを組み合わせるといろいろな知見が得られます。ですから、広域に分散した大量のRDFデータを統合的に利用できるデータベース環境は、情報の高度利用には不可欠です。グリッドの分野でも、応用の高度化に伴い、セマンティックWeb技術の適用が重要となりつつあり、従来扱っていなかったRDFデータを大量に保持・処理することが重要な課題となっています。

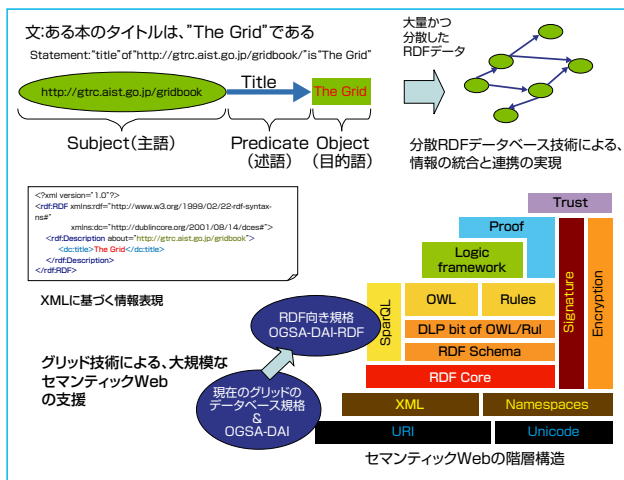
OGSA-DAI-RDF: 分散RDFデータベースのためのグリッドのミドルウェア

私たちは、こうした「セマンティック・グリッド」と呼ばれる分野で分散データベースを実現する基盤として、OGSA-DAI-RDFというソフトウェアを研究開発し、公開しています。これは、グリッドの標準アーキテクチャOGSA(Open Grid Services Architecture)に基づいたミドルウェアで、異なったデー

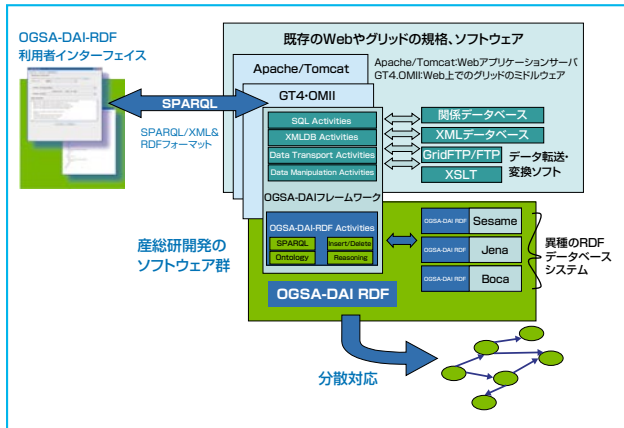
タベースソフトウェアで管理されているRDFのデータに対し、SPARQLという標準言語を使ってグリッドの環境から容易に、かつ統一的に検索できるようになるものです。グリッド上では、大規模な分散計算ソフトウェアや、異なるモデルの分散データベースとも連携できるので、グリッドの応用の高度化や大規模なセマンティックWeb応用に対するグリッド技術による効果的な支援が可能になりました。

規格化と国際的な開発との連携

私たちは、研究開発と並行して、OGF(Open Grid Forum)のデータベース作業グループ(DAIS-WG)に対して、RDFのアクセス規格を加えることを提案しています。規格を実現するソフトになることで、使える標準の実現を目指しています。



▲ RDFとそのデータ例



▲ OGSA-DAI-RDFの基本構成図