

D-029

SuperSQL を利用したデータベースシステムの構築と現状の問題点 Construction of Database System based on SuperSQL

五嶋 研人†
Kento Goto

鈴木 秀男‡
Hideo Suzuki

1. はじめに

データベースを用いた Web アプリケーションを作成する際、通常、HTML 内に PHP などのコードを書いてデータベースとのやり取りを行う。したがって、Web アプリケーション作成者は、HTML だけではなく、PHP や SQL にも精通している必要がある。逆に、これらを理解するために時間が取られて、主とするアプリケーションが完成するまでに、時間を要することもある。それに対して、SuperSQL と呼ばれる SQL の拡張言語を用いると、SQL に似た短いクエリを実行するだけでデータベースからデータを取り出し、それらを成型した形の HTML が作成される。本論文では、その SuperSQL の有効性を中心に、データベースとして使用した時の性能評価についても言及する。

2. SuperSQL について

SuperSQL とは、慶應義塾大学の遠山先生の研究室で研究・開発が行われている SQL の拡張言語で、関係データベースの出力結果を構造化し、多様なレイアウト表現が可能な言語である。SuperSQL のクエリは、SQL に似た文法体系となっており、以下の<media>の部分に出力したい媒体を指定することで、様々な媒体出力が可能である。

SuperSQLのクエリ構造

```
GENERATE <media> <structure> FROM <table>;
```

<media> : HTML,PDF,Excel,XML,etc.
<structure> : 出力構造
<table> : データベースのテーブル名

本論文では、<media>の部分に「HTML」を指定し、SuperSQL を用いた Web アプリケーション開発の有効性について検証した。

† 職業能力開発総合大学校東京校 応用課程 生産情報システム技術科

‡ 東海職業能力開発大学校 電子情報技術科

3. SuperSQL を用いたデータベースシステムの構築

本論文では、「数式表示・検索システム」を SuperSQL を用いて Linux 上に構築した。このシステムは、データベースに登録されている数式を、複数の形式で Web 上に表示させることが出来る。

本システムにおける処理の流れは、以下の図 1 の通りである。まず、ユーザーからのリクエストを基に、Web サーバが SuperSQL を介してデータベースへ問い合わせを行う。そして、その結果を SuperSQL が HTML として Web サーバへ返し、Web サーバは、その HTML をブラウザへ表示させる。

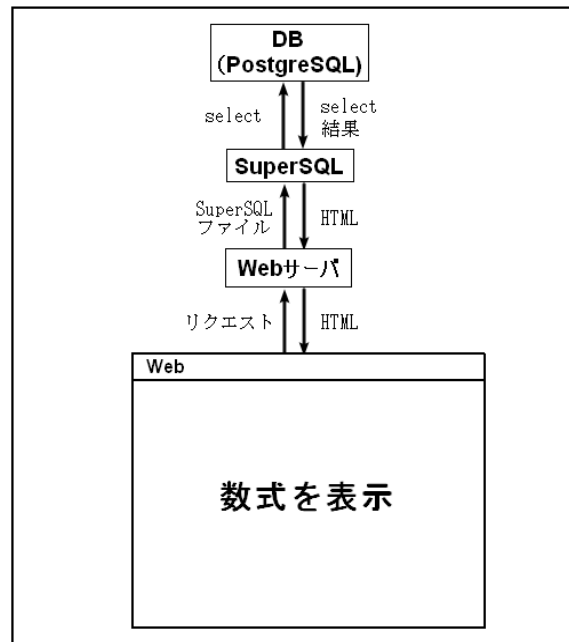


図 1 処理の流れ

また、このシステムの主な機能は、以下の通りである。

- ・数式を、LaTeX, MathML, 画像の 3 つの形式で表示
- ・登録されている数式を検索 (キーワード検索可)
- ・セッション管理を行い、管理者がデータの insert, update, delete を Web 上で行う

4. 利点と現状の問題点

4.1 利点

実際に、SuperSQL を使用して「数式表示・検索システム」を作成した際の利点を以下にまとめておく。

(1) クエリが覚えやすい

SuperSQL のクエリは、SQL に似た文法体系となっている。そのため、SQL に知見のある人にとっては、覚え易い。

(2) 短いコードで HTML を生成できる

HTML の作成に SuperSQL を用いると、通常の HTML 内に PHP を書いてデータベースへ接続する方法と比べて、遥かに少ないコードでデータベースの内容を表示させる HTML を生成することができる。また、Ruby on Rails と比較しても、SuperSQL を用いた場合の方がより短いコードで HTML を生成することが可能である。

(3) 開発効率のアップ

上記(1),(2)により、開発者は長いコードを書く必要がなくなり、また、PHP 等を新たに習得する必要もなくなり、総じて開発効率のアップに繋がる。

4.2 問題点

実際に、SuperSQL を使用して「数式表示・検索システム」を作成した際の問題点を以下にまとめておく。

(1) インストールが難しい

現在公開されている SuperSQL はパッケージ化されておらず、また、詳しい導入マニュアルも存在しないため、相当な知識を持った開発者でないとインストールは困難である。

今後 SuperSQL を普及させるためにも、早いうちにパッケージ化を行うことが望ましい。また、本研究の過程で作成した詳しいインストールマニュアルを、近いうちに公開する予定である。

(2) EUC-JP 準拠

慶應義塾大学理工学部の DBMS(PostgreSQL)が EUC-JP である都合上、SuperSQL は EUC-JP に準拠するよう開発されており、他の文字コードを考慮していない。そのため、データベースが EUC-JP 以外だった場合、またコンピュータの他の文字コードだった場合、うまく動作しない場合がある。

デフォルトの文字コードが EUC-JP 以外の OS も多数存在するため、ユーザーが実行時、またはインストール時に文字コードを選択できるようにすることが望ましい。筆者は UTF-8 でも使用できるように、研究室から入手した SuperSQL のプログラムを少し書き換えてコンパイルし直して使用している。

(3) 複数テーブルを参照すると、実行に時間を要する可能性がある

SuperSQL は、複数テーブルを参照したクエリを実行する際、各テーブルのデータ量によっては実行に時間がかかる場合がある。

例えば、5 個のテーブル a, b, c, d, e (格納データ数 a:20, b:100, c:10, d:10, e:5) に対して、それぞれの格納データ一覧を 1 つのページ (HTML) に表示させる 1 つのクエリを、コンピュータ (Intel(R) Core(TM)2 Duo CPU E6850 @ 3.00GHz, 2.00GB RAM) で実行してみたところ、実行時間は、「329,683 ミリ秒」であった。これは、現状の SuperSQL が、FROM 句の後ろに書かれている各テーブルの行数の直積分のデータを一度生成してから、各テーブル毎の処理を行う仕様になっているためである。例えば、上記の例の場合だと、 $20 \times 100 \times 10 \times 10 \times 5 = 1,000,000$ 行のデータが一度生成されてから処理が行われる。そのため、HTML 生成には非常に時間を要してしまう。

この問題に対する修正版の SuperSQL は、現在、慶應義塾大学の研究室において開発中である。開発はほぼ完了しており、そちらを用いると、上記の例の場合でも、1,000 ミリ秒以下の実行時間で HTML を出力することが可能である。

5. おわりに

4.1 の利点を見ても分かる通り、SuperSQL を用いると、より短いコードで Web アプリケーションの開発を行うことが出来る。このことにより、確実に開発効率は向上すると考えられる。

今後は、本論文で開発した「数式表示・検索システム」を発展させて、SuperSQL を用いた数式処理連携システムを開発しようと考えている。このシステムでは、数式処理システムとの連携も行い、結果を自動計算し、MathML, LaTeX, 画像の 3 つの形式で出力させる予定である。また、フォーム画面を用いて、ユーザーも数式を登録できるようなシステムにしたいと考えている。

謝辞

本論文の研究にご協力頂いた慶應義塾大学の遠山教授、及び遠山研究室の研究生の方々に、感謝の意を表します。

参考文献等

- [1] SuperSQL Web Page
<http://ssql.db.ics.keio.ac.jp/index.html>
 (引用)