Z39.50/SQL+ – Stateful Web Access to Relational Databases

Robert M. Colomb Sonya M. Finnigan
Distributed Database Unit, CRC for Distributed Systems Technology Centre
Department of Computer Science, The University of Queensland, Qld 4072, Australia
Email: s.finnigan@dstc.edu.au

Abstract

The ANSI/NISO Z39.50 Standard defines a protocol to facilitate the interconnection of computer systems for the search and retrieval of information in databases. In this paper, we present Z39.50/SQL+, the adaptation of this protocol to the SQL domain, and briefly discuss its advantages in terms of information retrieval.

Open SQL Environment

Structured organizational databases, typified by SQL databases at present, are rarely available on the public networks. Instead, their networks are closed, either limited to an organizational environment or proprietary products. They network in two ways, client-server and multidatabase, via middleware. In both cases, knowledge of system catalogues and query language dialects must be hardwired into the application or the middleware platform.

The technical problems of making an SQL database available on an open network environment are analogous to those of making a text database available on that environment. Firstly, the server must have a standard way of making its system catalogues available to the client, of agreeing on the query language dialects and types of data available, and a standard means to verify version compatibility between client and server. Secondly, the query language must support the ability of the server to not only process a query but to retain it as the basis for further queries, in the same sort of way as managing saved result sets in text databases. Finally, the server must support a standard means to manage and account for software actions (such as triggers) originating at a client but residing at a server site. All of these facilities are available in the Z39.50 protocol; the Z39.50/SQL+ project presents an adaptation of that protocol to an open SQL environment.

Z39.50/SQL+

Z39.50/SQL+ can be seen as an extension of the existing Z39.50-1995 (Version3) protocol, uniting the advantages of the SQL RDBMS's with those of Z39.50. This SQL extension is most beneficial (but not restricted to) when a client wishes to retrieve information from an SQL database.

The existing ANSI/NISO Z39.50-1995 communications protocol is an open standard designed to facilitate the search and retrieval of information in databases. The standard specifies formats and procedures governing the exchange of messages between a client and server, enabling a client to request that the server search a database and identify records which meet specified criteria, and to retrieve some or all of the identified records. It includes features which allow the server to advise the client as to the names and characteristics of the elements of the server database and to establish agreement as to what query language is in common between the client and the server. It provides features to manage state during a session, and to manage and account for state across sessions. Resource and access control facilities are also available.

Z39.50/SQL+ introduces a new query type, a type-SQL2 query - a query conforming to the SQL-92 standard which is highly structured allowing search terms and attributes to be specified within the query. Pre-defined attribute sets, which provide a virtual global data schema, are not necessarily required as the SQL server database already stores its metadata within its system catalogues. Like Z39.50, Z39.50/SQL+ still distinguishes two types of response records that may occur from the server: database and diagnostic records. It introduces a new record syntax, SQL2-RS, by which database records may be returned and similarly an additional error format, SQL2-ERR. No changes to the resource report and access control formats are envisaged at this stage. Minor extensions to the Explain record syntax include version and catalogue table name parameters.

Z39.50/SQL+ provides to the SQL user a stateful communication environment with the full flexibility and query power of SQL when connecting their working environment to a remote database. Z39.50/SQL+ clients will be able to formulate complex queries, either by using SQL or one of its derivatives, such as Query-by-Example (QBE). Queries may be formulated on multiple tables supporting cartesian products, unions, intersections, joins on matching columns, and projections on given columns, as well as being able to use powerful constructs for expressing conditions, performing aggregate and comparison operations, partitioning tables into groups and much more. In addition, SQL RDBMS's provide structured, organizational databases complete with data management facilities including system catalogues, flexible indexing and query optimization - providing efficient access and retrieval of both data and metadata.

Project Status

A full technical report, *'Z39.50/SQL+ Project'*, has been written and is available on the web (http://www.dstc.edu.au/DBU/research_news/z3950.html). The building of the prototype, using the YAZ Z39.50 toolkit, began in March. The project is scheduled to have a base-line implementation ready for demonstration in early October, with Explain and Extended Service facilities available in early '97.

In Summary

Z39.50/SQL+ presents an open SQL standard to facilitate stateful internet access for controlled information retrieval from remote relational databases. It is the facility to manage and account for state within and across communication sessions that distinguishes Z39.50/SQL+ from both existing SQL communication standards and proprietary middleware.

The work in this paper has been funded in part by the Cooperative Research Centres Program through the Department of the Prime Minister and Cabinet of the Commonwealth Government of Australia.