

SPDB Native XML Implementation Status

Evan Rosson

March 13, 2008

SPDB's native XML implementation is progressing on schedule. Notable work completed during winter quarter 2008 includes:

1. ExistDB Implementation Completed
2. Automated JUnit Tests
3. Right Join/Right Mix Implemented
4. Major Build Process Changes
5. Various Refactorings and Improvements
6. Progress on Experiments and Writeup

1 ExistDB Implementation Completed

The quarter began with only a proof of concept: simple selection and conditionalization were available from the SPDB server, not in ExistDB, but in another XML database; and a partial set of SPDB XQuery functions was completed. The SPDB server work was completed using ExistDB; and all XQuery functions supporting selection operations were completed. Database updates are mostly complete, as well: create relation, drop relation, and insert data are fully implemented. Delete data will be completed shortly.

2 Automated JUnit Tests

Numerous automated JUnit tests were added alongside the ExistDB implementation. Each new XQuery function is unit tested. Conversion of SPOQL statements to XQuery statements is also unit tested. Some (less thorough) unit tests were created to show Oracle's correctness in selected areas, as well.

3 Right Join/Right Mix Implemented

Right join and right mix are described in other SPDB papers, but did not exist in the Oracle implementation. This functionality was added to ExistDB. (Oracle remains unchanged.)

The following syntax is accepted in ExistDB:

```
SELECT [variables] FROM [relation] LEFTJOIN [relation]...  
SELECT [variables] FROM [relation] RIGHTJOIN [relation]...
```

Old syntax is still supported. Left join is assumed.

```
SELECT [variables] FROM [relation] JOIN [relation]...
```

4 Major Build Process Changes

The project is now built using Maven, rather than an Ant script. This has several benefits:

- JUnits are integrated with the build process.
- Easier to add new external dependencies to the project.
- Cleaner project directory structure.

More information on Maven is available at <http://maven.apache.org>.

Other build process changes:

- The documentation for the build process/project setup was improved.
- The SPDB server and clients are now packaged and run as .jar files.
- Client and server code is cleanly separated: SPDB clients can be built and distributed separately from the server.

5 Various Refactorings and Improvements

A short list of other substantial, but mostly transparent, changes:

- The processed query format - created by the SPOQL/Algebra parser and passed to the database adapter - was dramatically modified to improve type safety, testability, and extensibility. Database adapters now use a visitor pattern to parse processed queries, rather than the previous chain of if-statements. (Changes to the Oracle implementation were kept to a minimum.)
- Database connections are now threadlocal, replacing the previous fixed-size connection pool. This simplifies implementation and results in noticeably quicker JUnits and server startup.
- Database connections are now visible only to database adapters. This improves encapsulation, and was required to allow the implementation of non-JDBC database connections such as ExistDB.
- Error handling has been somewhat improved: some code now uses exception handling to deal with errors, rather than null values or string-based error handling.

6 Progress on Experiments and Writeup

Planning for experiments began this quarter. Important factors for experiments were identified, and existing experiments from the existing version of SPDB were examined. Experiments should be completed and runnable early in spring quarter 2008.

Experiments and writeup are the major tasks to be completed during spring quarter 2008.