

NewsForge

The Online Newspaper for Linux and Open Source

<http://programming.newsforge.com/>



Title Penn State's QFilter puts security in XML database queries

Date 2005.01.12 7:03

Author Jay Lyman

Topic

<http://programming.newsforge.com/article.pl?sid=05/01/06/188254>

New software developed at Penn State University promises to protect XML database queries and filter out unauthorized requests, thereby boosting query performance as many as 100 times, researchers said.

QFilter, created by PSU School of Information assistant professor Dongwon Lee and three others at the university, can bypass the typical access control modules built into individual databases for security. The software can be deployed with off-the-shelf databases without requiring substantial changes to them, according to Lee.

"In XML access control, many proposed solutions are available recently," he said. "But they are not practical in our opinion since they usually assume that underlying XML database has some kind of built-in security features. This is usually not true -- there is no known XML database product with security features, not to mention that there are not many XML database products themselves."

To make matters worse, Lee explained, when object relational databases such as DB2, Oracle, or SQL Server are used as the underlying database engine, the proposed solutions become useless because they use an XML model, as opposed to the relational model of the underlying database.

QFilter, Lee said, can sit between the users and the database and filter out unauthorized requests for data before a database responds to a query. The shift from data filtering to query filtering provides a practical solution to access control issues and also boosts query-response time by rejecting unauthorized requests earlier, Lee said.

Although QFilter is not yet in its final version, Lee said he is working with school officials to release the software in open source form next year.

"I'm currently discussing this with the Penn State IP office while my student is polishing the software," Lee said. "It may take awhile -- sometime in 2005, I hope. If ready, then we probably will go by either Berkeley [license] style or GNU style, but I'm not sure at this point."

A second lens on stream processing

Developed on Windows XP and written with Java and Galax -- an open source XML kernel developed at Bell Labs -- QFilter may take stream processing to a new level, Lee said. The process whereby a stream of data is viewed through the lens of query is reinforced by a second lens through another query with QFilter, Lee explained.

"The current implementation supports most of Xpath query language," Lee said. "In the future, we want to add more features to the list."

The Penn State assistant professor and the rest of the QFilter team -- PSU doctoral student Bo Luo, assistant professor Peng Liu, and associate professor Wang-Chien Lee -- presented a paper on the software in early November at the ACM Conference on Information and Knowledge Management in Washington D.C. Lee said the paper, "QFilter: Fine-Grained Run-Time XML Access Control with NFA-based Query Rewriting," was well received. The software, however, is still in its alpha stage, despite the authors having begun work in late 2002.

More efficient query response

There are other technologies to restrict access to databases, such as the view-based approach that creates different data views for each user. Since user credentials do not have to be checked once views are created, this can increase query response speed. However, as the number of users requesting access

grows or if views need to be frequently updated, this method can cause maintenance and storage headaches, QFilter researchers contend.

By rejecting unauthorized queries early on, QFilter can improve query processing time dramatically, depending on different queries and data, Lee said.

"When users ask invalid queries, QFilter detects and rejects them outright," he said. "Thus, it saves a lot of unnecessary database query processing."

When users ask valid queries, QFilter detects that the query is authorized using non-deterministic finite automata (NFA) -- which store access control policies and monitor query flows -- and informs the database. In turn, the underlying database does not require a security check, leaving it free to focus on regular query processing, Lee said.

When users send some queries that request both unauthorized and authorized data, QFilter "prunes out" the unauthorized request part upfront, Lee said.

"As a side effect, usually this pruning results in more optimized query processing," he said.

According to the Penn State team's experimentation, end-to-end processing time from when query was issued to when answers were returned improved between 10 and 100 times with QFilter, depending on the XML query and data types, Lee said.



© Copyright 2005 - NewsForge, All Rights Reserved

printed from [NewsForge](#), [Penn State's QFilter puts security in XML database queries](#) on 2005-01-12 14:34:57