

JACKIE'S Forum

SQL Query Engine



Jackie Jansen

With very little fanfare IBM introduced a new SQL query optimizer for the iSeries in V5R2. Although the existing query optimizer had served us well in the world of online transaction processing, the evolution of more web-based queries and analytical queries created demand for a new optimizer. As has happened many times in the past, as we introduce something new, we now need a name for something existing. The iSeries now supports the new SQL Query Engine (SQE) and the original Classic Query Engine (CQE).

When SQE was first shipped in V5R2 only the very basic SQL single table queries were directed to this optimizer. Adding the latest V5R2 database group PTFs and PTF SI07650 increased the functionality of SQE. One major new enhancement was support for basic join queries. With V5R3 more and more SQL queries are directed to SQE. Some examples of queries that are still directed to CQE include logical file references, using LIKE predicates, and specifying ALWCPYDTA(*NO).

Non-SQL interfaces such as Query/400, OPNQRYP and QQQQry will always be directed to CQE. As SQE evolves the intent is that all SQL based queries, such as Open Database Connectivity (ODBC) and Java Database Connectivity (JDBC) queries, will be directed to SQE.

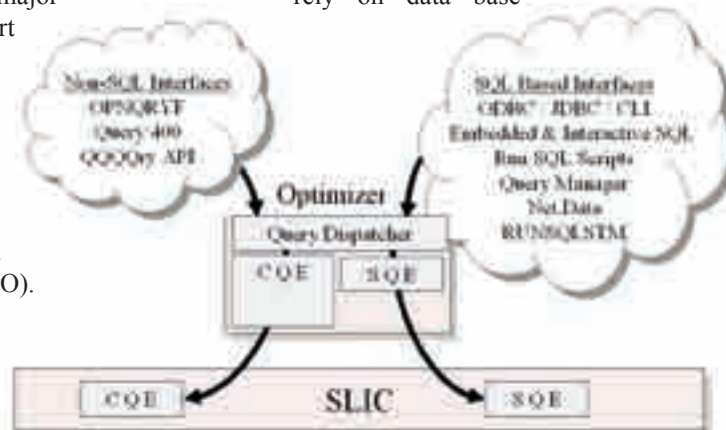
SQE can be very helpful in improving the performance of long running analytical queries. This is particularly appreciated in a Business Intelligence environment.

Two new components for DB2 UDB for iSeries include a Query Dispatcher and a Statistics Manager. The Query Dispatcher analyzes a query as it is

submitted and based on its attributes directs the query to either the CQE or the SQE query engines.

Prior to SQE the query optimizer performed the statistics retrieval. The optimizer used the table header to retrieve information such as total row count and table size. It also used existing binary radix indexes and Encoded Vector Indexes (EVIs) to help analyze the data in the underlying table. With V5R2, for SQE only, Statistics Manager provides an additional statistics source. The Statistics Manager is designed to collect column statistics and provide data to the optimizer. The Statistics Manager will collect statistics such as the number of unique values in a column (cardinality) and the most frequent values that occur in a column. Most databases today


rely on data base



administrators to collect and maintain their database statistics. By default on the iSeries, all statistics gathering is done automatically. When you first run a query that would benefit from statistics over a specific column, the system starts gathering the information for that column in the background. The next time you run the query the system will use the newly collected statistics and possibly modify the access plan. The system keeps track of how “fresh” the statistics are. Statistics are considered “stale”

if the number of rows in the table has increased by more than 15% since the statistics were gathered or if the number of rows that have been changed are more than 15% of the total row count for the table. If a statistic is referenced and it is “stale” the system will automatically start a job in the background to update the column’s statistics.

In a Business Intelligence environment it is very common to change a large number of rows at once during a data load. To help performance when you next run a query you can manually ask the system to update the column statistics ahead of time. For more details on this and many other features of SQE see the IBM SQE redbook SG24-6598.

You can use iSeries Navigator to view existing column statistics. Simply right click on the table name and choose “Statistic Data”. You can then view the “Details” of the individual column statistics. For more information on SQE see the IBM web site www.iseries.ibm.com/db2/sqe.html. 

Jackie Jansen is a Senior Consulting IT Specialist. She currently works in the IBM Americas Advanced Technical Support Solutions Centre. Jackie is a frequent speaker at iSeries Technical Conferences and User Group meetings. Contact her at jjansen@ca.ibm.com.