

Q&A Session for Chat with the Lab: "Improving Performance of Analytical Models with DB2 MQTs"

Session number: 669269415
Date: November 14, 2008
Starting time: 09:36 am

Willie favero - 10:05 am

Q: Is this based on LUW or DB2 in general

A: This presentation is based on LUW environment and DB2 LUW tools. However the MQT concepts presented are valid for DB2 in general.

Mark Leatherman - 10:21 am

Q: Is there any way to prevent logging during an MQT refresh? We have some very large MQT's and they require a large active log directory to handle the refresh.

A: Refresh defered - Full refresh

Avoid Logging Activity on the MQT by using one of the following:

- **ALTER TABLE ... NOT LOGGED INITIALLY**
On the same unit of work of the REFRESH statement
 - **Performing a LOAD instead of using the REFRESH statement**
A **LOAD FROM CURSOR** using the result of the MQT Query
Or **Load from a previously EXPORTED file** containing the result of MQT query
-

Kevin McCauley - 10:26 am

Q: Will MQT Matching occur without RI?

A: Yes. RI is defined for the base tables to improve MQT matching capabilities when the join in the MQT is not the same as the join in the query. It's not mandatory to define RI but when you do it DB2 compiler knows that the join between the tables with RI will not delete any rows due to a missing key (lossless join).

For Example: A sales_fact table with sales by customer and a cust_dim dimension table.

Let's define the following MQT:

```
SELECT f.cust_key, d.cust_name, SUM(f.sales)
FROM cust_dim d, sales_fact f
WHERE d.cust_key = f.cust_key
GROUP BY f.cust_key, d.cust_name;
```

Let's analyze the MQT matching with the following query:

```
SELECT cust_key, SUM(sales)
FROM sales_fact
GROUP BY cust_key;
```

With RI

Defining RI for these two tables (sales_fact, cust_dim) guarantees that you only have sales for existing customers and a query selecting only from sales_fact will have the same number of rows in the result set as a query that joins sales_fact table to cust_dim. (lossless join)
DB2 compiler will then match the MQT rerouting the query to use the MQT.

Without RI

Not having RI means that you could have sales for customer not present in the cust_dim table. In this case a query selecting only from sales_fact will have more rows in the result set than a query that joins sales_fact to cust_dim.
DB2 compiler will not match the MQT and will not use the MQT. DB2 will solve the query without the MQT.

You can have RI enforced by DB2 or not enforced by DB2.

Enforced: DB2 guarantees the RI. If the application tries to insert a row into sales_fact table with a cust_key that doesn't exist in cust_dim DB2 gives an error.

Not Enforced: The application is responsible to maintain the RI. DB2 will not block any violation to RI. So if you declare RI not enforced it is the user/application responsibility to guarantee it. This option tells DB2 compiler to consider the join as a lossless join because the application guarantees the RI.

Somasekar Srinivasan - 10:27 am

Q: why would there be a need for "disable query optimization" ?

A: You can temporarily avoid queries to route to a MQT for maintenance purpose. Another reason is to avoid the regular queries to route to the MQT and write specific queries using the MQT.

I B - 10:27 am

Q: Can we use refresh age <> ANY yet?

A: No. Deferred refreshed MQT will be considered only for refresh age = ANY.

Kent Rubin - 10:27 am

Q: MQTs are created as Tables in the database...I assume these can be exposed to Cognos for reporting from, correct?

A: That is correct. Cognos can see the MQT as a regular table. In that case you will model your query explicitly using the MQT.

Another option is to model the queries in Cognos using the base tables and not the MQT. In this case DB2 optimizer will match the queries generated by Cognos to the MQTs.

Kent Rubin - 10:28 am

Q: Or are there rules for using MQTs for reporting from?

A: No special rules are required for Cognos. The MQT matching happens inside DB2.

Javier Mazzurco - 10:28 am

Q: The CREATE TABLE can include LEFT JOINS?

A: Yes. You can have MQT with left outer joins but no query will be matched to this MQT. You will need to write the query using the MQT.

Krishna Venkatrama - 10:29 am

Q: Are MQTs considered a short term solution for creating performance datasets? against ETL best practices to create tables instead of MQTs for long term ?

A: You can create MQTs as performance databases and have the MQTs refreshed by the ETL processes. The advantage of MQT is the matching mechanism. You don't need to rewrite your queries to use the MQT. DB2 will do it automatically.

Yogi Dwivedi - 10:31 am

Q: Why would we loose data without RI?

**A: This is the definition of an inner join.
The rows in the right table with key values that do not exist in the left table will not be part of the result set.
The rows in the left table with key values that do not exist in the right table will not be part of the result set.**

With RI it never happens. RI guarantees that every key value in the right table is present in the left table and every key value in the left table is present in the right table.

ellen reys - 10:33 am

Q: will you discuss connection between MQT and Cubing Services? What additional advantage will CUBing Services provide?

A: Cubing Services stores the cubes in memory. When the OLAP application issues a MDX query against cubing services it answers with the data in memory. If the

data is not in memory Cubing Services queries the database (SQL) to cache the necessary data. When it happens, DB2 will try to match the query (SQL) from Cubing Services to an existing MQT.

tom lavelle - 10:45 am

Q: Since query patroler is being replaced with WLM, how are you interfacing with WLM

A: You can create and enable event monitor for activity and enable collection of activities using the COLLECT ACTIVITYE DATA ... WITH DETAILS clause on the CREATE or ALTER WORKLOAD STATEMENT.

Example:

```
CREATE EVENT MONITOR DB2ACTIVITIES FOR ACTIVITIES WRITE TO TABLE
SET EVENT MONITOR DB2ACTIVITIES STATE 1
ALTER WORKLOAD SYSDEFAULTUSERWORKLOAD
    COLLECT ACTIVITY DATA ON COORDINATOR WITH DETAILS
```

You can apply DB2 advisor to the queries collected by the DB2ACTIVITIES monitor:

```
db2advis -d SAMPLE -type IM -wlm DB2ACTIVITIES
```

You can query the SQL statements from ADVISE_WORKLOAD table.

```
Select statement_text from advise_workload
```

Krishna Venkatrama - 11:00 am

Q: Can we range partition MQTs ?

A: Yes. I recommend below link for a better understanding:

<http://publib.boulder.ibm.com/infocenter/db2luw/v9r5/index.jsp?topic=/com.ibm.db2.luw.admin.partition.doc/doc/c0021562.html>

Yogi Dwivedi - 11:03 am

Q: Followup question: the original query will be returning a row...why not MQT return the same set of rows?

A: For Example: A sales_fact table with sales by customer and a cust_dim dimension table.

Let's define the following MQT:

```
SELECT f.cust_key, d.cust_name, SUM(f.sales)
FROM cust_dim d, sales_fact f
```

```
WHERE d.cust_key = f.cust_key
GROUP BY f.cust_key, d.cust_name;
```

Let's analyze the MQT matching with the following query:

```
SELECT cust_key, SUM(sales)
FROM sales_fact
GROUP BY cust_key;
```

With RI

Defining RI for these two tables (sales_fact, cust_dim) guarantees that you only have sales for existing customers and a query selecting only from sales_fact will have the same number of rows in the result set as a query that joins sales_fact table to cust_dim. (lossless join)

DB2 compiler will then match the MQT rerouting the query to use the MQT.

Without RI

Not having RI means that you could have sales for customer not present in the cust_dim table. In this case a query selecting only from sales_fact will have more rows in the result set than a query that joins sales_fact to cust_dim.

DB2 compiler will not match the MQT and will not use the MQT. DB2 will solve the query without the MQT.

Barani Govindarajan - 11:05 am

Q: How does defining a MQT work on HADR databases?

A: MQT are synchronized as any other table on HADR databases.

Sandra Tucker - 11:05 am

Q: Will the optimizer choose MQTs for multi fact table join? so if there is a MQT for each of the fact tables in a query and has all the join points can it use the MQTs for each part and join the results?

Calisto Zuzarte - 11:22 am

A: Hi Sandra, we can take this one offline as the answer might be a little more involved. calisto@ca.ibm.com