Optimization Task Guidelines

SQL Server Profiler allows you to capture and analyze MDX statements performed by the SSAS server. As a result, it is perfectly suited for detecting slowly executed and incorrectly written queries.

To get started with the task, first you have to fill the data warehouse with significant amount of data. Ideally, you should have backup of your DW (file with the .bak extension) and load it into MS SQL Server. Secondly, you should open your implementation project in SSDT for Visual Studio, add the restored database as a Data Source, click Data Source View designer (where you can see all the tables in grey), right-click any blank space and select “Refresh”. Data source view should now be connected to the right data source. Now, process the cube.

Working with SQL Server Profiler

1. Open SQL Server Profiler.

2. Choose New Trace from File menu.

3. The tracking file can be created on the basis of one of two templates: the default Standard template allows you to capture information about the execution of all the most important MDX instructions, the Replay template allows you to capture all the information needed to retrieve the intercepted instructions later. We will define the tracking file manually:

a. In the Name field, enter “Test for ”

b. Expand the list Use the template and select the Blank template.

! Remember to connect to the Analysis Services Server type !

4. The collected data can be saved to a file or to a table. We will save the data in the file. Enter the file name and select Save.

5. Go to the Event Selection tab. It allows you to select the intercepted events and specify what information about each of these events will be stored in the trace file. Events are divided into different categories. From our point of view, the Query Processing category is particularly important, in which there are events reported during MDX execution, such as reading data from the cube or dimension, reading data from the buffer or formatting the results. Pay special attention to the Query Cube End event, which shows the query execution time in milliseconds (Duration).

6. Select some of the events you are interested in.

7. The number of events captured can be further reduced by means of filters. To limit them to events reported to one analytical database:

   1. Click the Column Filters button.

   2. Select the DatabaseName.

   3. Expand the Like condition and enter the name of the analytical database for your database.

8. Start the tracking by clicking Run and perform various operations on your data warehouse. In particular, make various queries (use Pivot Table in Excel to perform them).

9. Return to SQL Server Profiler, stop the trace file, and review the data collected in it.

Measuring the processing time of a cube using SQL Server Profiler

To capture the time it took the cube to be processed one needs to select the "Command End" event from the "Command Events" category. The minimal selection of columns for this event is:

- Duration
- Event Subclass
- Text Data

To locate the cube processing event look for events with a "12 - Batch" subclass. There are two commands of this subclass typically issued to the analytical database: deploy and process.

Those two can be distinguished by their Text Data. Deploy command contains an XML document describing the full structure of an analytical database, which is quite long. Proces command is much shorter and will contain the phrase "ProcessFull". Make sure not to use data for "ProcessDefault", which sometimes can also be issued.

An example text data for process full command:

```
<Batch xmlns="http://schemas.microsoft.com/analysisservices/2003/engine">
  <Parallel>
    <Process xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:ddl2="http://schemas.microsoft.com/analysisservices/2003/engine/2" xmlns:ddl2_2="http://schemas.microsoft.com
/analysisservices/2003/engine/2/2" xmlns:ddl100_100="http://schemas.microsoft.com/analysisservices/2008/engine
/100/100" xmlns:ddl200="http://schemas.microsoft.com/analysisservices/2010/engine/200"
xmlns:ddl200_200="http://schemas.microsoft.com/analysisservices/2010/engine/200/200"
xmlns:ddl300="http://schemas.microsoft.com/analysisservices/2011/engine/300"
xmlns:ddl300_300="http://schemas.microsoft.com/analysisservices/2011/engine/300/300"
xmlns:ddl400="http://schemas.microsoft.com/analysisservices/2012/engine/400"
xmlns:ddl400_400="http://schemas.microsoft.com/analysisservices/2012/engine/400/400"
xmlns:ddl500="http://schemas.microsoft.com/analysisservices/2013/engine/500"
xmlns:ddl500_500="http://schemas.microsoft.com/analysisservices/2013/engine/500/500">
      <Object>
        <DatabaseID>BookstoreDW2</DatabaseID>
      </Object>
      <Type>ProcessFull</Type>
      <WriteBackTableCreation>UseExisting</WriteBackTableCreation>
    </Process>
  </Parallel>
</Batch>
```

Creating physical cube structure

The physical structure of the cube is created on the basis of the groups of measures belonging to it. As a rule, each measure group reads data from another fact table. The SSAS server allows you to select for each measurement group one of three physical models: MOLAP, ROLAP or HOLAP.

The MOLAP model

In the default MOLAP model, all the data of the measure group is located in the analytical database, i.e. it contains a copy of the fact table and all aggregations calculated during the processing of the cube. The MOLAP model offers the highest browsing performance of cubes (because it does not require connection to the data warehouse and read the information it needs) and reliability (no connection to an external data source will prevent the user from viewing the cube). The disadvantage of this model is the storage of duplicate data in the analytical database - as a result, the analytical database and data warehouse store the same large amounts of data, and changes to the source data are not visible when viewing the cube. Processing causes the cube to be synchronized with the changed source data.

The ROLAP model

In the ROLAP model, all measure group data is stored in a data warehouse. In this model, both the fact table and the aggregates calculated for a group of measures are stored in the relational database, where the aggregations are stored in the indexed views created by the SSAS server, and thus require that the server has appropriate permissions on the data warehouse side. The main disadvantage of the ROLAP model is the low performance of viewing and processing cubes: if the data displayed is not in the SSAS server buffer or the client application, they must be read from an external data source. In addition, relational servers generate and read aggregations much slower than OLAP servers.

The HOLAP model

The HOLAP model is an indirect model - only aggregations are found in the analytical database, the content of the fact table is not copied to it. As a result, viewing detailed data in the analytical cube is slower because it requires reading from the warehouse. In this model, only the analysis of grouped data is fast, because the SSAS server reads aggregations from the analytical database. The advantage of the HOLAP model is the lack of storing copies of the same data in two databases (warehouse and analytical base).

In each of the three models, the structure of the analytical cube (metadata forming them) is stored in the analytical database.

	MOLAP	HOLAP	ROLAP
Querying time	Short	Moderate (short with well designed aggregations)	Long
Processing time	Long	Moderate (if no aggregations are designed, it will be short)	Short
Total size	Big (size of the measure group is much smaller if no aggregations are designed for them)	Moderate	Small

Analytical cubes and measure groups are logical objects, while their physical representations are partitions. Each measure group must contain at least one partition. The first partition is created automatically when you add a group of measures to the cube.

To change the physical data model for partitions for a group of measures, in the BIDS project for your data warehouse:

1 the Partitions tab.

2. Expand the partitions for the measure group of choice.

3. Select Storage Settings and set a different model type (MOLAP is the default model for all partitions).

4. Process the cube.

5. For different models, enable tracing in the SQL Server Profile tool and compare cube processing times and query execution times.

To check the estimated total cube size right click on the database in SQLSMS and select Properties. One of the properties is *Estimated Size*.

Defining aggregations – example for the bookstore DW

Aggregations are calculated during the processing of cubes, results defined for individual measures of grouping functions. The SSAS server calculates the aggregates so quickly when performing queries that in the case of small (less than one million rows) fact tables, their lack will be unnoticeable to the user.

To design the aggregates, open a data warehouse project in the BIDS tool and:

1. Go to the Aggregations tab.

2. The main window will display both groups of cube measures for bookstore - select Sprzedaz Ksiazki measurement group.

3. There is a Design Aggregation button on the left side of the toolbar - click it.

4. After clicking Next you will be asked to change the default aggregation level defined for each dimension attribute. Let's assume that users do not group data by book titles, but regularly group them by genre - change the level of aggregation for the attribute Title to None and for the attribute Genre to Full.

5. After clicking Next you will be asked to enter the number of values of particular measures and attributes. If you won't do this, the wizard will design statistically the best aggregations, so entering incorrect data will result in designing suboptimal aggregations. Click the Count button.

6. The next question of the wizard will concern the method of aggregation design.

a. By selecting Estimated storage reach, you can enter the maximum size of the partition, after which further aggregations will no longer be generated. This option should be chosen if the size of the analytical database is the most important factor.

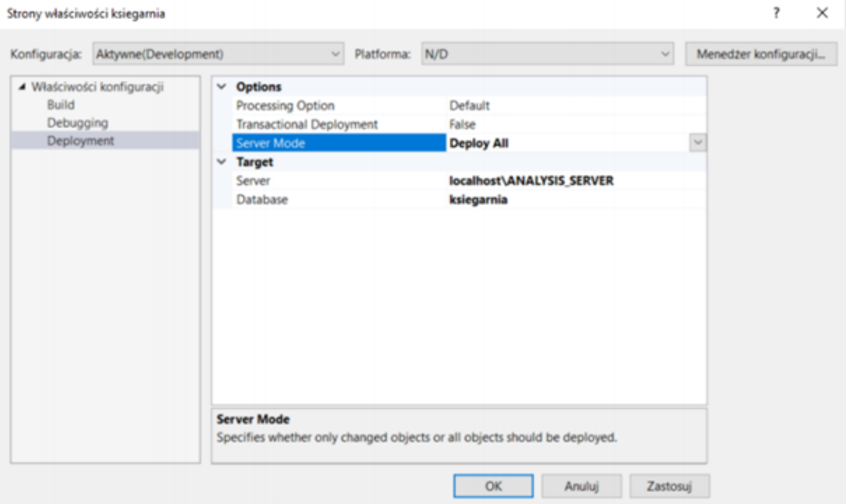
b. By choosing the option Performance gain reach, you can enter the average, percentage reduction in the time of viewing the analytical cube. In this case, you should start designing the aggregation from the default, 30% performance improvement, and then adjust the level of aggregation based on server performance monitoring. Choose this option.

7. Click Next and select Deploy and process now.

8. Run the SQL Server Profiler tool and compare query execution times for analytical databases with designed and not designed aggregations for different physical models (MOLAP, ROLAP and HOLAP).

IMPORTANT NOTES:

● You should set “Deploy All” option on the server mode before cube processing



Strony właściwości ksiegarnia

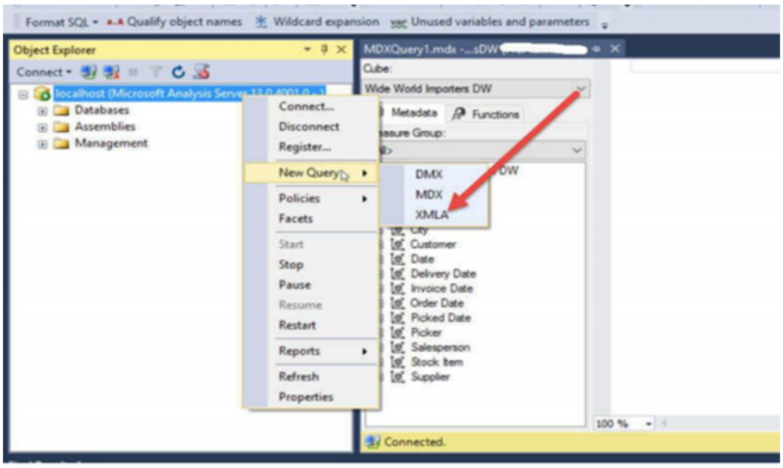
Konfiguracja: Aktywne(Development)

Platforma: N/D

Options	
Processing Option	Default
Transactional Deployment	False
Server Mode	**Deploy All**
Target	
Server	**localhost\ANALYSIS_SERVER**
Database	**ksiegarnia**

Server Mode
Specifies whether only changed objects or all objects should be deployed.

● You should clear cache before every query execution



```
<ClearCache xmlns=http://schemas.microsoft.com/analysisservices/2003/engine>
   <Object>
      <DatabaseID>your analytical database ID</DatabaseID>
   </Object>
</ClearCache>
```

To get your analytical database ID right click the database --> Properties and see the “ID” attribute.

• **Duration of query executions should be the mean of at least 10 executions of the same query.**

GOOD LUCK!

Ostatnia modyfikacja: poniedziałek, 22.05.2023, 13:46

Optimization Instruction

Przejdź do...

Optimization report (template) ►

eNauczanie to usługa elektroniczna oparta na popularnym systemie Moodle, umożliwiająca szeroko rozumiane kształcenie zdalne, odbywanie zaliczeń testów weryfikujących wiedzę oraz dostęp do wirtualnych laboratoriów. Nowoczesna platforma dydaktyczna oferuje obecnie blisko 5000 kursów z różnych dziedzin wiedzy, obsługując ponad 20 tysięcy użytkowników rocznie.

Prywatność | Regulamin platformy | Deklaracja dostępności

Szablon e-kursu do przedmiotu

Portal informacyjny eNauczanie PG

Pomoc eNauczanie

Kalendarz roku akademickiego

Eduroam (sieć bezprzewodowa)

Do końca sesji pozostało: 15:33:31

Pobierz aplikację mobilną

Tw ja wygaśnie za **5 minut (11:26:10)**. Zapisz dotychczasową pracę!