Jan Dzięcielski 191702

Kamil Lesiński 189501

***Optimization of the Data Warehouse (Car Sharing)***

***Aim of the laboratory***

The aim of this laboratory is to observe and discuss issues concerning various physical cube models and aggregation design.

***Preliminary assumptions***

Size of the database:

3984 MB

Testing environment:

* Windows 10 Home x64
* Risen 7 4800HS
* RAM 16GB
* Microsoft SQL Server Profiler 2019
* Microsoft Visual Studio 2022
* Microsoft SQL Server Management Studio 2019

***Testing***

|  | | MOLAP | | ROLAP | | HOLAP | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Aggr. | No aggr. | Aggr. | No aggr. | Aggr. | No aggr. |
| Querying speed (for 3 different queries) | query 1.4 | 15 ms | 17 ms | - | 83 ms | 75 ms | 75 ms |
| query 2.7 | 117 ms | 126 ms | - | 178 ms | 172 ms | 179 ms |
| query 1.2 | 5 ms | 17 ms | - | 174 ms | 4 ms | 171 ms |
| Processing time | | 5603 ms | 3223 ms | - | 1 912 ms | 4308 ms | 4387 ms |
| Total size | | 52,04 MB | 52,05 MB | 21,66 MB | 21,80 MB | 21,78 MB | 21,80 MB |

***Discussion***

***Querying time***

Querying time for MOLAP is comparatively the fastest, since this model does not need connection to the data warehouse - it stores the copy of all the needed data. There is not that much difference in executing times with the use of aggregations, the only significant difference is with executing query 1.2, for which aggregation speeds up this process.

ROLAP has a long querying time due to reading data from external data source.

For the HOLAP model, querying time is comparable to ROLAP times, the difference being in computing query 1.2. The HOLAP model uses aggregations from analytical databases, which shortens the time of completion of this query.

***Processing time***

Processing time in the MOLAP is significantly higher than in ROLAP and slightly higher than in HOLAP. It is due to the aforementioned copy which is contained in MOLAP, because the processing is synchronized with the changed source data. HOLAP processing time is slightly lower, because it is indirect, which means it only contains aggregations in analytical database. On the other hand, ROLAP processing time is much quicker due to having indexed views of the fact table and aggregations stored in it.

***Total size***

Total size in the ROLAP and HOLAP models are similar - about 22 MB, whereas the size of the MOLAP is the biggest - 52MB. The MOLAP model takes that much space because the analytical database contains a copy of the fact table and all aggregations calculated during the processing of the cube. The reason for lower size of the ROLAP and HOLAP model is that for the ROLAP model all measure group data is stored in a data warehouse, and in HOLAP, only aggregations are in the analytical database.

***Important note***

ROLAP cannot be checked for aggregations due to error considering creation of indexes on the fact table aggregation view, it cannot reference derived table.