dimensio informatics

"AI instead of KIWI"

# Performance Boost for Siebel with Artificial Intelligence

**The mightiness of Oracle Siebel ("the world's most complete CRM solution") may also be the source of various performance related issues. The application's flexibility is based on a highly dynamic, generic data model which allows Siebel users to define their own business objects, a feature that may cause problems with regard to query optimization. Even without any individual enhancements, Siebel queries are often very complex and push the optimizer of the database management system to its limits. Response times in the range of minutes or even hours leave users disappointed. The attempt to address performance issues with conventional tuning methods often fails and faster hardware (keyword: "kill it with iron") may only marginally improve the situation.**

**The way out is called "Artificial Intelligence". We offer an intelligent solution with the multi-dimensional semantic DIMENSIO index that is based on self-learning neural networks. This enables us to accelerate slow queries by three orders of magnitude (1,000-fold) without modifying the application. The software intervenes—completely transparent to Siebel—and enhances complex SQL queries with additional information which helps the DBMS optimizer to generate better execution plans. The application example of a Swiss insurance company demonstrates how Siebel queries can be sped up from hours to only seconds (see reverse page: "Die Mobiliar").**

## Long response times - long faces

The majority of Siebel queries result in joins of many tables. The crux, however, is that the restrictions on the individual tables are not particularly selective, i.e. the DBMS has to materialize a number of joins one after the other before being able to reduce the intermediate result to the relevant result set. This leads to read operations in the gigabyte range on the disks or in the buffer cache as well as large amounts of data that need to be moved around in order to return just a few hits in the end. Experience has shown: database analyses such as a segmentation that take hours to be processed or supposedly simple queries which have a response time of several minutes, inevitably lead to user frustration. In a worst-case scenario, this can lead to poor user acceptance of certain functions or even to a boycott of the entire system.

## Conventional tuning methods don't work

As the SQL queries are generated dynamically at runtime due to Siebel's generic data model, the DBA has no other choice but to resort to conventional tuning methods: creating indexes and providing statistical data. Indexes, however, do not only offer advantages: they also take up a lot of storage space and the overhead for data modifications increases considerably. Moreover a large variety of different arrangements of indexed columns would have to be created to support a wide range of query scenarios. In terms of performance, statistical data are also somewhat critical. In case they are outdated, a good execution plan can fail und response times may take longer without any discernible reason.

## "KIWI" is not an option either

If conventional tuning methods have reached their limits, the only other option to address performance issues is to stock up on "iron". Additional CPUs, however, can relax the situation if the processors (or cores) are the limiting factor and if the DBMS can execute the queries in parallel. In this case more Oracle database licenses for additional cores would have to be purchased. It is usually the read operations that use up most of the resources. More main memory or faster disks might help in this case. It is true that more RAM can be utilized to increase the size of the database cache but this is of little use

**dimensio informatics GmbH • Brueckenstrasse 4 • 09111 Chemnitz/Germany • Phone: +49 371 262019-0**
**Fax: +49 371 262019-10 • info@dimensio-informatics.com • www.dimensio-informatics.com**

if the DBMS has to scan entire tables or to materialize joins in the first place. And even flash memory-based solid state drives (SSDs) offer only little help: they do accelerate disk accesses but they do not solve the basic problem of too many read operations.

## The solution: Multi-dimensional semantic indexing with AI

Performance improvement measures can only be successful if they tackle the problem that complex queries impose restrictions on many attributes in dozens of tables. Traditional database indexes are one-dimensional and can only map columns of exactly one table. They are not helpful for complex queries where each index alone reduces the number of hits in a single table only to a limited extent and only the combination of several attributes leads to an adequate selection.

The magic formula is called "multi-dimensional indexing" of any number of columns (regardless of their order) in any number of tables. Originally classic multi-dimensional indexes were developed for geocoded data in a two-dimensional space. However, they structure the data room according to purely formal criteria, i.e. without knowing the data semantics and are limited to only a few dimensions. This formal limitation can be overcome by "knowing" the semantic data relationships—a task that can be achieved by methods using "artifical intelligence".

The patent-pending, multi-dimensional semantic DIMENSIO index is based on a self-learning process employing neural networks. During the learning phase the knowledge for building the index structure is acquired. This includes the analysis of SQL queries and the identification of relevant attributes which will restrict the result set. In the next step the index is built which is able to map thousands of dimensions. For this purpose queries are run against the database to extract the primary keys of the tables in question. The attribute and key values found are then handed over to the self-learning process. This machine learning method clusters data according to their similarities in a "semantic map" and stores them in the DIMENSIO index. This is then used to optimize the database queries—completely transparent for the application.

The trick is: Besides knowing the attributes used as criteria for selecting data, the index also knows the primary key values of the relevant database tables. Each Siebel query is now being analyzed and run against the DIMENSIO index. The result is a number of hits in form of primary key values. The original SQL query is then supplemented by this list of values, providing the DBMS optimizer with an extremely selective query. It will therefore choose an execution plan which uses the primary indexes thus requiring only a few cache or disk accesses. This translates into sometimes drastically reduced response times of the Siebel system. An extreme example is a "customer segmentation" with 9 result rows which DIMENSIO was able to reduce from 3.5 hours to 2 seconds.

The AI-based process adjusts to changing data sets. Newly acquired knowledge about added records or changed values complements the existing index without having to examine the analyzed data again.

## Minimally invasive approach

DIMENSIO functions transparently as well as in a minimally invasive way. The application software remains untouched and only minor adjustments are required for the synchronization service of the database. In Siebel environments the application server is chosen as the integration point with the indexing software running on its own server. During runtime the DIMENSIO software operates between the communication of the application server and the DBMS. It analyzes the SQL queries and evaluates the potential benefit which can be achieved with the DIMENSIO index. If the potential is high, the query is enhanced with data from the index. If no benefit is to be expected from the index, the query will be passed through without any change. The overhead for the analysis lies in the nanoseconds range and can hardly be measured.

## Conclusion

The DIMENSIO software can easily be integrated into Siebel environments. The implementation process usually takes only a few weeks starting with the first analysis, through the proof of concept and finally to going live. During operation the software runs silently and requires only a minimum of administration and maintenance. The sustained benefit sets in immediately: an immense SQL acceleration which boosts Siebel queries into the range of perceived "real-time".

## Reference customer: Die Mobiliar

At the **Swiss Mobiliar Insurance Company**, around 80 general agencies care for more than 1.2 million customers. The implemented Siebel CRM solution was no longer able to cope with the administration of customer and contractual data, nor could it support the marketing activities. Long response times were the consequence. DIMENSIO succeeded in reducing the duration of the SQL queries from more than 3 hours to 2 seconds, which translates into an acceleration by a factor of about 6,000. At the same time, the server load of 26 GB was reduced to 5 MB – which corresponds to a reduction by a factor of about 6,300.

*„Up until now, the usage of the CRM system was limited by the enormous amount of data to be processed.*
*With DIMENSIO, data retrieval was tremendously boosted thus contributing significantly to an increased efficiency in data handling."*

(You will find the complete case study on our website: www.dimensio-informatics.com)