

# Technical Introduction to the IBM Smart Analytics Optimizer for DB2 for System z

Namik Hrlle, Oliver Draese

IBM Germany - Research and Development GmbH  
Schoenaicherstrasse 220  
71032 Boeblingen, Germany  
{hrlle,draese}@de.ibm.com

The IBM Smart Analytics Optimizer for DB2 for z/OS is a new technology to extend existing data warehouse environments on IBM mainframe systems. It is a workload optimized appliance that enables customers to analyze huge amounts of data in a matter of seconds instead of minutes or hours by delivering unmatched performance. This doesn't only allow “train-of-thought” analysis as interactive scenario but also enables business requests which were simply impossible before. Analytical workloads can now be executed as a online process instead of asynchronous batch processing. A call center employee can for example analyze the customer's behavior pattern while he still is on the phone.

To achieve this new performance, the Smart Analytics Optimizer is implemented as a distributed, In-Memory system where a cluster of computing nodes holds the data in a specialized format in main memory structures. New technology enables the product to perform scans over compressed data without the need of decompression prior to applying predicates. A special partitioning scheme allows the parallel processing of the data with as few locking mechanisms as possible. As the industry trend is showing that an increase of single thread performance is no longer achievable but even standard computers are now delivered with multiple CPU cores, the Smart Analytics Optimizer is designed to exploit this new hardware as good as possible by assigning specific subsets of data to specific cores. The product by itself is running on a cluster where standard instances own hundreds of cores and terabytes of main memory. But even within a single computing core, the product makes use of SIMD instructions to perform parallel evaluation of predicates on multiple tuples.

Besides the raw performance of this new product, the deep integration might even be considered more important. The Smart Analytics Optimizer is not a stand-alone product as it is offered by several other vendors. Instead it extends the existing relational database manager (DB2) by its functionality without requiring any changes to the existing application environments.

Programs, which were connecting to DB2 before just continue to execute their workload against the mainframe database. The internal DB2 functionality then decides when to make use of the new Smart Analytics Optimizer or not. The granularity for these decision is a query block. This implies that a single query with multiple query blocks can be partially executed on the Smart Analytics Optimizer and partially on the mainframe directly. The joined results are returned back to the requesting application by DB2, hiding the complexity of the different execution environments and the required transformations.

Trademarks: IBM logos are trademarks or registered trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Multiple DB2 subsystems can share a single Smart Analytics Optimizer to make optimal use of the attached hardware. Maintenance and configuration is done from DB2 side too. The new product is handled just like any other DB2 resource. Stored procedures and new DB2 commands are used to query the status, perform configuration tasks or control the resources of the accelerator. One of the main goals is to simplify the work of the database administrator. Instead of adding a new resource that needs to be managed, the Smart Analytics Optimizer is an appliance that controls and maintains most of its functionality by itself. The general database administration is now easier because the administrator now longer has to try anticipating all the different query requests against existing tables to create the optimal index structures. By attaching the accelerator, the queries that would fail to match an index, resulting in an expensive table scan, are now routed to the Smart Analytics Optimizer and still answered in a matter of seconds.

Our presentation will give a technical overview about the Smart Analytics Optimizer, typical workloads and the new approaches how these are handled. It explains the new techniques from a hardware and software perspective.