

New opportunities to drive analytics value into business operations

Use the IBM DB2 Analytics Accelerator



Contents

- 3 Accessing enterprise data in place
- 4 Gaining advocates from IT, application teams and Lines of Business
- 4 Uncovering and expanding opportunities for the DB2 Analytics Accelerator
 - Accelerator-only tables
 - DB2 Temporal Data Management
 - IBM Netezza In-Database Analytics
 - DB2 Analytics Accelerator Loader for z/OS
- 8 Case studies
- 10 Measuring the business value of the DB2 Analytics Accelerator
- 11 The potential for the DB2 Analytics Accelerator to provide even greater ROI
- 11 Conclusion

Today, many professionals are using the IBM® DB2® Analytics Accelerator (the Accelerator) to help their organizations gain even greater insight and value from their data. The Accelerator is a high-performance appliance for DB2 for z/OS®. Organizations can offload data-intensive and complex DB2 for z/OS queries to the Accelerator in order to support data warehousing, business intelligence and analytic workloads. The Accelerator executes these queries quickly, without requiring CPU utilization by DB2 for z/OS.

The Accelerator is a logical extension of DB2 for z/OS, so DB2 manages and regulates all access to the Accelerator. DB2 for z/OS directly processes relevant workloads, such as OLTP queries and operational analytics. Queries that run more efficiently in a massively parallel processing (MPP) environment are seamlessly rerouted by DB2 for z/OS to the Accelerator. There is one set of credentials that is governed by RACF® security, and all access flows through DB2 for z/OS.

Users often first see the business value of the Accelerator in handling long-running queries, but many are also finding that the Accelerator can drive cost savings in areas such as administration, storage and consolidation as well as delivering real-time analytics.

This white paper discusses how organizations can improve analytic insight with the IBM DB2 Analytics Accelerator. It offers guidance to help organizations more quickly uncover new opportunity areas where the Accelerator can have the greatest impact. The paper covers topic areas including:

- Accessing enterprise data in place
- Gaining advocates from IT, application teams and Lines of Business
- Uncovering and expanding opportunities for the DB2 Analytics Accelerator
- Measuring the business value of the DB2 Analytics Accelerator
- Case studies
- The potential for the DB2 Analytics Accelerator to provide even greater ROI

Accessing enterprise data in place

For many enterprise organizations, data originates on IBM z Systems® infrastructure. This is where they run their mission-critical applications in order to take advantage of industry-leading qualities of service such as availability, reliability and Evaluation Assurance Level (EAL) 5+ security. Many organizations then replicate or transfer that data — including sensitive data — off of z Systems. This is typically done in order to realize perceived cost savings, to combine z Systems data with data from external sources, or to satisfy end-user technology preferences.

However, this approach can create security and data governance risks. As organizations copy or ETL data to disparate platforms managed by separate people, processes and infrastructures, the risk and liability associated with a data breach dramatically increases. In addition, Lines of Business often support their own data stewardship policies that can conflict with corporate data governance guidelines. For example, each group may have its own policies for encryption, its own separate set of credentials, its own third-party access policies, its own regulatory compliance policies and more. Moving data between platforms can also amplify complexity, increase batch windows and elevate costs. The Accelerator can help organizations avoid these risks by providing access to enterprise data in place while making it available to analytics applications in real time.

The Accelerator and compliance

For more information on the Accelerator and compliance (security and regulatory) read the white paper: [Is your business risking a USD 100 million loss? How new technologies can help protect your data and your bottom line.](#)

For more information on the rising costs associated with the liability of a security breach read the [Ponemon Institute 2016 Cost of Data Breach Study: Impact of Business Continuity Management.](#)

Gaining advocates from IT, application teams and Lines of Business

To validate the impact that the DB2 Analytics Accelerator can have on your business, IT applications and infrastructure teams need to communicate closely with the Lines of Business. The Lines of Business leaders whose teams use the applications can help IT uncover areas of the business that can benefit most from the Accelerator and quantify the return on investment (ROI).

IBM offers guidance to help IT leaders facilitate this team approach through a [Cost Benefits Analysis \(CBA\) workshop](#). This workshop can help you begin to socialize the value of the Accelerator with application owners and business users, uncover use cases throughout the organization and start to validate ROI.

“... no one has ever been able to speak to our application teams about technology, let alone build a business case with them. What you guys were able to do, working with them to build a case with this technology is just amazing.”

— IBM CBA workshop client

Uncovering and expanding opportunities for the DB2 Analytics Accelerator

There are four main areas where clients are gaining value from introducing the Accelerator into their environment:

- **Rapid acceleration of business-critical queries**
Long-running and complex queries historically run on DB2 for z/OS now run dramatically faster on the Accelerator.
- **Reduced IT sprawl for analytics**
The Accelerator can simplify and consolidate complex infrastructures and reduce time-consuming ETL processes, minimizing the need to offload data to a distributed system. This can help reduce data latency, enhance reliability, increase security and decrease overall TCO.
- **Business insight derived from z/OS transaction systems**
The Accelerator provides a hybrid transactional and analytics environment giving users real-time access to enterprise data on z Systems.
- **Improved access to historical data and lower storage costs**
The Accelerator enables organizations to cost-effectively store large volumes of historical data. That data remains accessible online so it can be seamlessly leveraged for analytic queries.

In addition, there are powerful new capabilities available in DB2 Analytics Accelerator Version 5 that add even greater value in driving insight to the business.

Accelerator-only tables

Version 5 introduces Accelerator-only tables. As their name suggests, the tables only exist in the Accelerator. This new capability helps organizations expand the business-critical queries that can be processed on the Accelerator.

Accelerator-only tables can be used to further simplify data-transformation processes by virtually eliminating the need for data marts and complex ETL processes. This can help your organization reduce IT sprawl associated with analytics. In addition, Accelerator-only tables allow organizations to derive real-time insight from z/OS transactional systems. This means you can now create a work area for data scientists by defining a set of Accelerator-only tables that can help provide deeper insight into customers and markets.

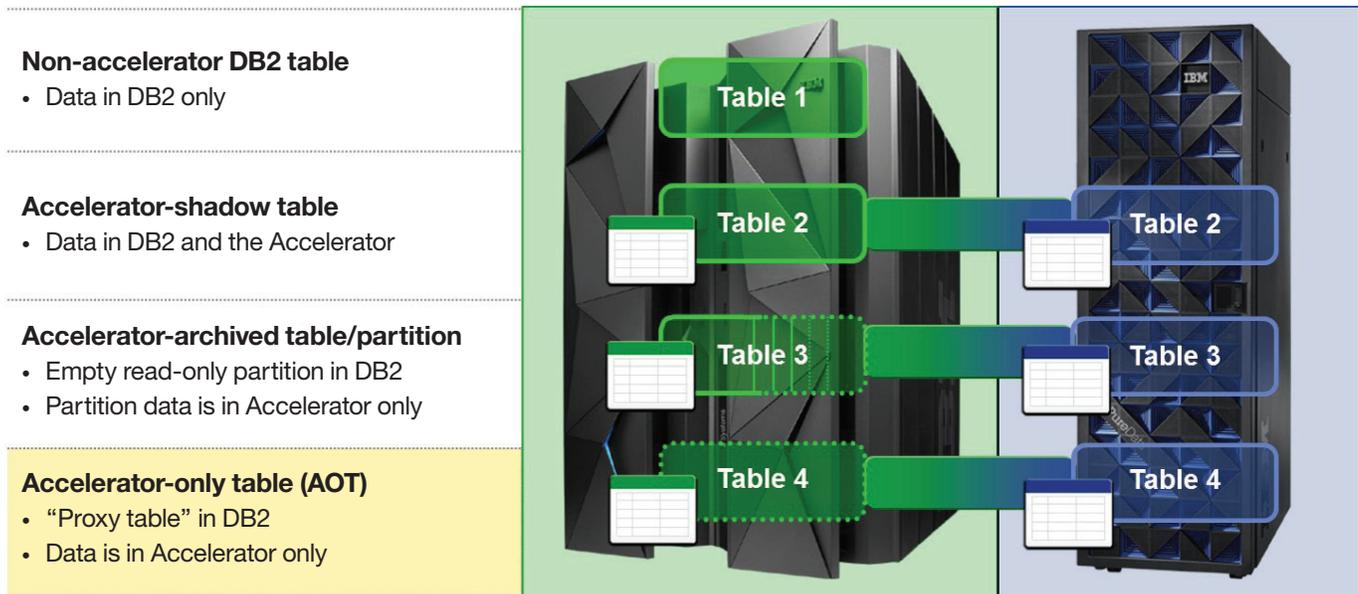


Figure 1: Accelerator-only table.

The DB2 Analytics Accelerator Loader for z/OS (Accelerator Loader) can now load non-DB2 for z/OS data into Accelerator-only tables. This simplifies data integration, while enhancing access to historical data and lowering storage costs. Accelerator-only tables can be used with IBM Campaign (formerly Unica® Campaign), QMF, or proprietary applications, as well as other reporting tools such as MicroStrategy. For example, DB2 Query Management Facility (QMF) can leverage Accelerator-only tables to save intermediate results temporarily on the DB2 Analytics Accelerator as part of a multi-step process and persist a query result in an Accelerator-only table for later accelerated processing. For more information, this [Redbook](#) provides additional detail about Accelerator-only tables and also includes load performance test results. Various runs show Accelerator-only table load speeds of 100 million rows in fewer than 30 seconds, 400 million rows in fewer than 50 seconds, and 4 billion rows in 500 seconds, with the entire load processing occurring on the DB2 Analytics Accelerator.

DB2 Temporal Data Management

Organizations are facing increased requirements to maintain historical data as well as a history of all data changes. Some of these requirements are based on government regulations where auditing and compliance inquiries require point-in-time analysis of data over any given time in the past 10 years. Many organizations have recognized the need to maintain a full change history in order to analyze trends, develop predictive models and perform statistical analysis such as regression analysis. They have created their own complex ETL processes, landing zones and data repositories simply to maintain a complete change history of data originating from DB2 for z/OS. This complexity can be greatly reduced by leveraging time-based (temporal) data management support within DB2 for z/OS and the Accelerator. DB2 for z/OS temporal support allows you to insert, update, delete, and query or snapshot data on demand—in the past, present and future—while keeping a complete history of what you knew and when you knew it.

When temporal versioning support is added on a DB2 for z/OS table, DB2 automatically tracks when changes (inserts/updates/deletes) are made to the state of a table, automatically preserves deleted records in a temporal table with period timestamps, automatically maintains period timestamps for inserted records, allows for partitioning and archiving of the temporal data, and makes the data available online for analysis.

The Accelerator also includes an online archiving functionality called the High Performance Storage Saver (HPSS). Using this functionality to move historical data into the Accelerator reduces the data volume in the DB2 for z/OS table; this also means less administration and fewer indexes. With Version 5, temporal tables can now be range partitioned, and the historical data can be archived to the Accelerator. This [Redbook](#) and the [IBM developerWorks® Community](#) discuss temporal support in DB2 for z/OS and the DB2 Analytics Accelerator in detail.

IBM Netezza In-Database Analytics

The IBM Netezza® In-Database Analytics capability is now available in DB2 Analytics Accelerator Version 5.1. This capability is targeted at users and developers interested in leveraging the development and use of analytic data modeling algorithms to perform research or other business-related activities. This support brings data mining capabilities to the DB2 Analytics Accelerator platform, enabling in-database data mining on large data sets using the computational power and parallelization mechanisms provided by the Netezza appliance. Most currently available data mining tools suffer significant performance limitations when applied to large data sets. Limited memory size and external storage requirements can make data modeling against large datasets impractical.

To overcome both these limitations, the parallel architecture of the Netezza database environment enables high-performance computation on large data sets, making it the ideal platform for large-scale data mining applications. Netezza In-Database Analytics support on the DB2 Analytics Accelerator consists of the following data modeling algorithms: Decision Tree, Regression Tree, Naïve Bayes, KMeans clustering and SPSS® TwoStep clustering. Through the invocation of DB2 for z/OS stored procedures, users can leverage in-database data modeling on the DB2 Analytics Accelerator, take advantage of the heavy integration of SPSS Modeler, and integrate these models with transactional applications running on DB2 for z/OS to improve the accuracy, speed and performance of real-time scoring, while reducing cost and complexity.

DB2 Analytics Accelerator Loader for z/OS

The DB2 Analytics Accelerator Loader for z/OS extends the analytic capabilities of the Accelerator with the ability to easily and efficiently bring non-DB2 data from other sources to the Accelerator and IBM z Systems servers. The combined use of the Accelerator, z Systems and the Accelerator Loader provides a single location with combined data sources such as IBM IMS™, VSAM, sequential files, ADABAS and Oracle to deliver faster insights for enhanced decision-making and lower storage costs. In addition, the Accelerator Loader includes built-in functions to load most of the System Management Facility (SMF) records and z/OS system log records to the Accelerator, which provides the capability to store this information in the Accelerator for system operation analysis.

The Accelerator Loader can provide the following benefits:

- In-memory transformation (ETL) of non-DB2 data, leading to significant cost and time reduction by virtually eliminating manual processes
- Insight into more data types such as IMS, VSAM, sequential files, Oracle, ADABAS, and more
- Reliable data availability for analytics applications
- Direct load of Accelerator-only tables for greater savings
- Data additions to existing table to avoid reloading entire table
- Support for consolidating changes made to replicated table or tables

Case studies

Large organizations in many industries have already realized significant business benefits from the Accelerator. These case studies provide some examples, but are not exhaustive representations of the Accelerator's functionality.

A major US insurance company speeds up delivery of analytics to the business

A major US-based insurance company has a significant amount of IMS data from legacy systems that comprise the bulk of their policy master data. They need to run analytics queries that join the data from both IMS and DB2 for z/OS to do policy quality scoring. They initially were only able to extract and load a random sample (approximately 10 percent) of their IMS data into DB2 for z/OS. The ETL process was taking up to two days to complete and the queries ran from 12–14 hours. As a result, this analytics process could only be run on a monthly basis on a small portion of the data. Since it was very CPU-intensive, it required changing Cobol extract and load programs any time the Lines of Business wanted to add or remove columns of data. Using the Accelerator Loader, the client is now able to load 100% of the policy data into Accelerator-only tables in a matter of two to three hours and the analytics queries run in under 45 minutes. The load process is zIIP-enabled, which reduces CPU usage. The Lines of Business can now run the analytics application weekly or more frequently if needed.

A large US company makes three years of data available online

Another organization wants to run operational analytics on their DB2 SMF data to improve performance problem identification and capacity planning. Their environment contains over 80 DB2 for z/OS production subsystems, so the data volumes are massive. They can only maintain approximately 14 days of detailed SMF data and 3 months of summarized SMF data. In addition, understanding SMF record layouts and formatting SMF data in order to load it into DB2 for z/OS can be a daunting task without additional tools. An IBM proof of concept for the Accelerator Loader that is currently in progress is showing excellent preliminary results. By using the Accelerator Loader with its predefined SMF record views, they are now able to quickly load all of the desired SMF records into the Accelerator. They estimate that they will be able to keep three years' worth of detailed data and make it readily available online in the Accelerator without consuming expensive direct-access storage. Queries and reports can run in the Accelerator to gain even greater mainframe CPU savings.

A large financial institution solves a big data challenge

A large financial institution is using the Accelerator Loader to process and store large amounts of interchange transaction data in the Accelerator. These interchange records were previously stored on tape and accessed by custom programs that were written to calculate fee adjustments and take other actions. Using the Accelerator, they can now use SQL to identify the transactions that need adjustments and produce the adjustment records—dramatically reducing complexity and processing time. This data was previously considered to be too voluminous to store in a relational database, but the high compression rates and large storage capacity of the Accelerator now make it possible.

A large investment firm reduces latency and costs in their analytics system

This organization was performing ETL to move DB2 z/OS data to a non-z/OS Oracle data warehouse. They were only able to perform ETL once per day, and were concerned with ETL costs, latency and poor performance. To address this, they implemented the same data warehouse using the Accelerator to significantly reduce the data latency, ETL costs and complexity. They also needed to bring non-DB2 data, including Oracle, VSAM and Tealeaf[®], into the solution.

Their original approach involved a three-step process:

1. Writing a custom program to extract data from those sources
2. Executing the DB2 Load utility to load this data into DB2 for z/OS
3. Initiating the load process from DB2 z/OS to the Accelerator

With the Accelerator Loader they can now use Dual Load when they need data in both DB2 for z/OS and the Accelerator and they can use the Accelerator alone when they do not need data in both. They use Load Resume in order to avoid the need to extract the VSAM, Oracle or Tealeaf data in its entirety. They can now extract only the new records and append those to both DB2 for z/OS and the Accelerator in parallel or append them to just the Accelerator-only table.

Measuring the business value of the DB2 Analytics Accelerator

The beginning of this white paper discusses the importance of socializing the Accelerator to the Lines of Business to uncover potential business areas that can benefit most from the technology. It is equally important to have a feedback system that can determine the effectiveness of the Accelerator and to be able to share best practices and lessons learned among

Lines of Business. Combining both technical feedback from monitors and SMF data along with feedback from the Lines of Business helps provide a clear picture of how the Accelerator is being used and how efficiently capacity is being managed.

DB2 monitoring tools, such as [IBM OMEGAMON® XE for DB2 Performance Expert](#) provide statistical and accounting information about use of the Accelerator. Metrics are available to monitor the Accelerator's operational status, query execution (per query or aggregated), statistics and accounting. There is also a set of tools on the [developerWorks](#) website that is designed to provide additional monitoring and data gathering capabilities. The links provided contain additional information on these tools and the data that is provided by each. Learn more at the [IBM Knowledge Center](#).

In addition to technical data provided by the monitoring tools, it is important to understand how the Lines of Business are using the Accelerator. There is no substitute for effective intra-office communication. This can help drive new use cases and ensure that the Accelerator support team is aware of new projects. It can help support teams ensure sufficient availability of resources and help them proactively plan for new appliances in order to meet ongoing demand.

The potential for the DB2 Analytics Accelerator to provide even greater ROI

Future releases will continue to increase the Accelerator's transparency to applications, enable more query acceleration and enhance existing capabilities. A Unified Store technology will strengthen Hybrid Transactional and Analytical Processing (HTAP) capabilities by virtually eliminating latency between data creation and data consumption. Delivering Accelerator functionality in the cloud will complement on-premise deployments, further speed up its return on investment, and add cloud value propositions such as on-demand provisioning and elasticity. The query engine will continue to evolve and be a leader in the market by permanently embracing the best technologies from IBM research labs and the industry such as BLU Acceleration®. The Accelerator will join BLU Everywhere with a new appliance that incorporates the best of Netezza and BLU Acceleration. A Hybrid Cloud Data Warehouse will be possible using the new appliance along with dashDB™ with BLU Acceleration and DB2 with BLU Acceleration using Webscale compute for seamless integration.

Conclusion

The DB2 Analytics Accelerator continues to drive new value into the business while extending current processes to do more. The key to recognizing and beginning to exploit possible use cases is to set up effective, results-oriented communication among IT, application teams and Lines of Business personnel. Once the organization-wide requirements are understood, the business cases for the Accelerator will become clear. IBM is ready to work with you to accomplish that goal.

About the authors

Shantan Kethireddy

Master Inventor, Sr. Certified IT Specialist,
Big Data & Analytics, IBM

Jim Knisley

Analytics CTP, Software Sales, IBM

Annette Zawacki

Portfolio Marketing Manager, Enterprise Analytics Strategy
and Execution Lead, IBM

Patricia Zakhar

Senior Marketing Manager, Analytics on z Systems, IBM



© Copyright IBM Corporation 2016

IBM Corporation
Route 100
Somers, NY 10589

Produced in the United States of America
October 2016

IBM, the IBM logo, ibm.com, DB2, z/OS, RACF, IBM z Systems, Unica, SPSS, IMS, Tealeaf, OMEGAMON, BLU Acceleration, and dashDB are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml.

Netezza® is a trademark or registered trademark of IBM International Group B.V., an IBM Company.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions. It is the user’s responsibility to evaluate and verify the operation of any other products or programs with IBM products and programs. **THE INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED,**

INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NONINFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation. Statements regarding IBM’s future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Statement of Good Security Practices: IT system security involves protecting systems and information through prevention, detection and response to improper access from within and outside your enterprise. Improper access can result in information being altered, destroyed, misappropriated or misused or can result in damage to or misuse of your systems, including for use in attacks on others. No IT system or product should be considered completely secure and no single product, service or security measure can be completely effective in preventing improper use or access. IBM systems, products and services are designed to be part of a lawful, comprehensive security approach, which will necessarily involve additional operational procedures, and may require other systems, products or services to be most effective. **IBM DOES NOT WARRANT THAT ANY SYSTEMS, PRODUCTS OR SERVICES ARE IMMUNE FROM, OR WILL MAKE YOUR ENTERPRISE IMMUNE FROM, THE MALICIOUS OR ILLEGAL CONDUCT OF ANY PARTY.**



Please Recycle