Understanding DB2 for i and the DB2 product family

Kent Milligan
ISV Business Strategy and Enablement
July 2010
**Table of contents**

Abstract........................................................................................................................................1

Introduction ..................................................................................................................................1

Distinguishing the three DB2 family members........................................................................2

  Understanding the benefits of shared development ...............................................................3

  Explaining portability concerns ..............................................................................................3

    DDL ...................................................................................................................................3

    DML ...................................................................................................................................4

    Syntax variances .................................................................................................................4

    Common programming interfaces .....................................................................................4

    Administrative controls ......................................................................................................5

    Query optimizer ..................................................................................................................5

    Middleware ..........................................................................................................................5

Summary......................................................................................................................................6

Resources....................................................................................................................................7

About the author ..........................................................................................................................8

Trademarks and special notices...............................................................................................9
Abstract

This white paper discusses the IBM DB2 family of products. The reader will learn about the history and the IBM goals for development, compatibility and standardization of these offerings. This white paper also focuses on the functional differences between the DB2 for i product and its two sister offerings.

Introduction

An integrated, fully relational database has shipped with every IBM® i (formerly IBM System i®) server as far back as the inception of the IBM AS/400® systems in the late 1980s. The DB2 branding for this integrated database was first adopted in 1995 with the product name, DB2/400. After several releases which advanced the functionality of the database, the IBM DB2 Universal Database™ branding was added and the product named changed to IBM DB2® UDB for AS/400. The product name was later changed to IBM DB2 UDB for iSeries® when the hardware platform was renamed. DB2 for i is the current product name. DB2 for i delivers advanced SQL functions and performance to customers and software developers. Although the product has evolved through name changes and enhancements, DB2 for i is still known for its integration and ease of use that have attracted and spoiled IBM i customers for years.
Distinguishing the three DB2 family members

Along with DB2 for i, there are two other members in the DB2 product family, as follows:

- For many years, DB2 was exclusively available on the mainframe. That product is now known as IBM DB2 for z/OS®.
- DB2 for i was the last database product to join the DB2 brand, but as mentioned, the core of its relational database engine was developed in the 1980s.
- The youngest database member is DB2 9.7 for distributed platforms – it is the single product that is available across all Linux®, UNIX® and Microsoft® Windows® platforms. This product is also known as IBM DB2 for Linux, UNIX and Windows. This product is actually the founding member of the DB2 Universal Database brand, but that branding was dropped in 2006.

The most frequent misconception about DB2 branding is this: It infers that IBM implemented a single, common code base on all of the supported platforms and operating systems. On the contrary, each DB2 brand member’s code version is unique and developed by different IBM laboratories, but a tremendous amount of technology sharing takes place at various levels across IBM and the DB2 brand. The different code bases allow IBM to exploit the hardware, microcode and operating system of each of the respective platforms. Thus, at the lowest level, each DB2 offering is tightly integrated into, and exploitive of, its operating environment. This is good news, as you can imagine, and as you will come to understand more clearly as you continue reading.

Another misconception is that the DB2 9.7 product is the master version of DB2 – meaning that the SQL functionality it contains is a superset of all the functionality in the DB2 product family. Much of that confusion is caused by the fact that the version numbers used by DB2 for i version numbers are completely independent from the version numbers used by the other two DB2 products. The DB2 for i version number matches the version of the IBM i operating system with which it ships. The reality is that each DB2 product has features that it delivered first to the market and that is unique to that product. For example, DB2 for i was the first DB2 product to support SQL stored procedure and is the only product that supports encoded vector indexes.

Furthermore, the various DB2 family members have differing approaches to scalability. The IBM i platform running IBM i has a philosophy based on scaling up, not scaling out. Because of this, the feature and function priorities of DB2 for i can be different than DB2 for LUW. For example, DB2 for i tends to have very high limits to growth, without the need to resort to partitioning and distributed queries. In contrast, DB2 for i does indeed support local table partitioning if required. In general, a scale-up architecture that provides excellent price-performance is much simpler to configure, manage and use as compared to a scale-out architecture. The DB2 for i scale-up approach is in line with the IBM i strategy for solving business problems as quickly and easily as possible.
Understanding the benefits of shared development

At the lowest level, the IBM DB2 development teams have ported code for a new database feature from another code version directly into their own. This sharing helps the DB2 brand members deliver new features to the market quicker.

At a higher level, IBM investments in pure research projects have resulted in many patented technologies that benefit all of the DB2 offerings. Of particular relevance are the number of patents in the field of query optimization and performance. In addition, the DB2 product family also has garnered a wealth of knowledge and experience from supporting mission-critical applications with relational database management systems since the inception of the relational architecture. These patented technologies and real-world experiences have enhanced reliability, performance and scalability across the entire DB2 product family.

For the DB2 developer and user community, the most important level of sharing and standardization occurs at the SQL language level. Each DB2 development team places a genuinely high priority on standardizing its product on a common SQL language. For example, if support for a new function that is being added to one DB2 product already exists in another, the SQL syntax for the existing version is reused.

Additionally, all DB2 development teams place heavy emphasis on delivering SQL syntax that conforms to the international SQL standard. The SQL 2003 standard is a good example. The IBM DB2 team provided input during creation of the SQL 2003 standard and implemented functions adhering to the standard as it was developed.

Explaining portability concerns

Having made the two previous points, the current level of SQL support found in the DB2 product line does not provide 100-percent portability of SQL across the DB2 brand. See the Resources section of this white paper for a web site that lists the common DB2 features. With the delivery of i5/OS V5R4 in early 2006, DB2 for i developers were proud to be the first DB2 family member to comply with the core level of the SQL 2003 standard. DB2 for i is also compliant with the core level of the most recent SQL standard — SQL 2008.

When looking at the issue of SQL portability, SQL statements break down into two types:

- Data Definition Language (DDL) for creating and managing database objects
- Data Manipulation Language (DML) for accessing and processing data stored in these objects

**DDL**

DDL usually requires more modification than DML when moving between the DB2 product members because of the differences in the underlying operating systems. A good example of this is that there is no SQL syntax support for table spaces in DB2 for i. This is because the IBM i operating system automates all low-level space allocation and management tasks. In fact, the DB2 for i V6R1 release added syntax tolerance to allow it to ignore SQL syntax associated with operating system and lower-level configuration setting that does not apply to DB2 for i.
DML

DML has a higher level of syntax compatibility (close to 95 percent) across the DB2 brand because operating-system differences tend to have minimal impact on the retrieval and modification of relational data. However, vestigial differences in SQL syntax persist, because many of these differences serve to exploit the target platform fully.

Syntax variances

Functional differences are more of an issue than syntax variances when porting to other DB2 product members. These differences arise from the fact that not all new features become available at the same time across all platforms. This difference in timing happens because of varying release schedules and unique user requirements that exist for each DB2 product. The recent enhancements in DB2 for i 7.1 and DB2 9.7 greatly improve the functional equivalency across the DB2 Database product line. See the Resources section of this white paper for a web site that lists the common DB2 family SQL features. The Resources section also lists the DB2 family SQL reference web site.

Common programming interfaces

SQL standardization across the DB2 product line provides a common set of application programming interfaces for database access. Each DB2 member supports the following database interfaces:

- Microsoft ODBC
- Microsoft ADO.NET
- SQL CLI
- Sun Microsystems™ JDBC
- SQLJ
- PHP ibm_db2 extension

In addition, each DB2 product also provides SQL precompilers that allow a developer to embed static and dynamic SQL in portable C and COBOL application programs.

Distributed Relational Database Architecture (DRDA), from The Open Group, defines a common set of flows and protocols for distributed SQL applications, enabling applications to avoid the need to be aware of the particular flavor of the targeted DB2 platform. Furthermore, IBM provides data connectivity and integration software, IBM DB2 Connect™ (based on DRDA), that makes access entirely transparent to applications and also masks the need to know the networking protocol in use. DB2 9.7, DB2 for z/OS and DB2 for i all support DRDA-based access. Support for the XA protocol for distributed transactions is also available across the DB2 members.

Previously, the stored procedures feature had a number of differences across the DB2 brand. DB2 for i was the first brand member to provide support for an SQL procedural language based on the PSM standard, and now the SQL procedure language is available across all of the DB2 products. In addition, all DB2 products support Java stored procedures.

The latest releases of DB2 and DB2 for i remove some key differences by delivering support for XML data type, XML publishing functions, enhanced SQL procedural language and the Merge statement. Some other DB2 brand features that are not yet available on the IBM i platform include: Label-based Access Control, DB2 Spatial Extenders, XQuery support and Regression aggregate functions.
Administrative controls

Another component with differing levels of function is administrative controls. These differences are the result of the unique administrative requirements of each underlying operating system. Several administrative functions are unavailable for DB2 for i because the database manager and operating system already handle them automatically. For instance, DB2 for i does not provide a RUNSTATS utility for optimizer statistics because its database manager keeps these statistics current at all times.

Likewise, there is no concept of table spaces. By default, DB2 for i does not support the notion of independent, isolated databases. Instead, IBM implements DB2 for i as a single systemwide database. For example, DRDA access on another DB2 server requires specifying a database name on the SQL CONNECT statement. With DB2 for i, this database name is, in reality, the system name; when connected, the client can access any database object to which it has authority. Schemas are the logical containers for related database objects on the IBM i platform. Any user with the proper authorization can access a database object within a schema (without a CONNECT statement).

In addition, database relationships such as a referential constraint are definable across objects that are stored in different schema. Because these differences exist, slightly different sets of skills and tools are needed to support the three DB2 product members. The IBM Data Studio and DB2 Control Center products have limited support for DB2 for i servers. Thus, it is best for administrators to use the IBM i Navigator tool (formerly, IBM Systems Navigator) for DB2 for i databases.

Query optimizer

Each DB2 product features an advanced, cost-based query optimizer. Because every optimizer fully exploits its target system and hardware, each database engine has a unique performance personality. For example, DB2 for LUW multidimensional clustered indexes (MDC) provide a specific benefit based on the underlying system’s architecture. The DB2 for i optimizer does not require or support this type of index because of the capabilities provided by IBM i single-level storage and data-access methods. Therefore, it might be necessary to do some tuning when applications are moved to another DB2 product to best exploit the underlying system. As mentioned, however, much patented optimization technology and algorithms are common across the different query optimizers.

Middleware

IBM provides a common set of middleware and tools that work with each of the DB2 product members. As explained, DB2 Connect delivers common DRDA requester middleware to use with DB2 servers for the IBM i and z/OS environments. IBM even provides a specific packaging of DB2 Connect for IBM i users called DB2 Connect Unlimited Edition for System i. The IBM Change Data Capture and IBM DataPropagator™ products deliver data-replication capabilities across the DB2 family. The IBM InfoSphere® Federation Server product is a sophisticated multiple data-source product that permits DB2 brand members to transparently access heterogeneous database servers, such as Oracle, Sybase and SQL Server. (See the Resources section of this white paper for the InfoSphere Federation Server web site.) DB2 Content Manager includes integrated workflow capabilities to manage images, documents, web content, multimedia assets and more. You can also use the IBM Data Studio to develop, deploy and debug Java and SQL stored procedures for any of the DB2 products. The IBM InfoSphere Data Architect is a collaborative data-design solution that you can use to discover and model databases residing on any DB2 server.

Understanding DB2 for i and the DB2 product family
Summary

Even though some differences exist, the DB2 product line provides a common technology base and tools for building applications. With time, IBM will make it easier for both users and software vendors to leverage their skills and application investments across the DB2 product family. In addition, both online transaction processing (OLTP) and business intelligence solutions benefit from the fact that DB2 distributed, DB2 for z/OS and DB2 for i are optimized to fully exploit their native platforms.
Resources

These web sites provide useful references to supplement the information contained in this document:

- IBM Systems Information Center

- DB2 for i online references
  ibm.com/systems/i/db2/books.html

- IBM Publications Center

- IBM Redbooks™
  ibm.com/redbooks

- DB2 Family Common Features Matrix & SQL Reference
  ibm.com/developerworks/db2/library/techarticle/db2common

- SQL reference for DB2 cross-platform development

- InfoSphere Federation Server
  ibm.com/software/data/infosphere/federation-server

- InfoSphere Data Architect
  ibm.com/software/data/optim/data-architect
About the author

Kent Milligan

Kent Milligan is a DB2 technology specialist in IBM ISV Solutions and Strategy Enablement. Kent spent many years of his IBM career as a member of the DB2 development group in Rochester, Minnesota. He speaks and writes regularly on various DB2 for i relational database topics.
Trademarks and special notices

© Copyright IBM Corporation 2010. All rights Reserved.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

IBM, the IBM logo and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

Microsoft, Windows, Windows NT and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product or service names may be trademarks or service marks of others.

Information is provided "AS IS" without warranty of any kind.

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.