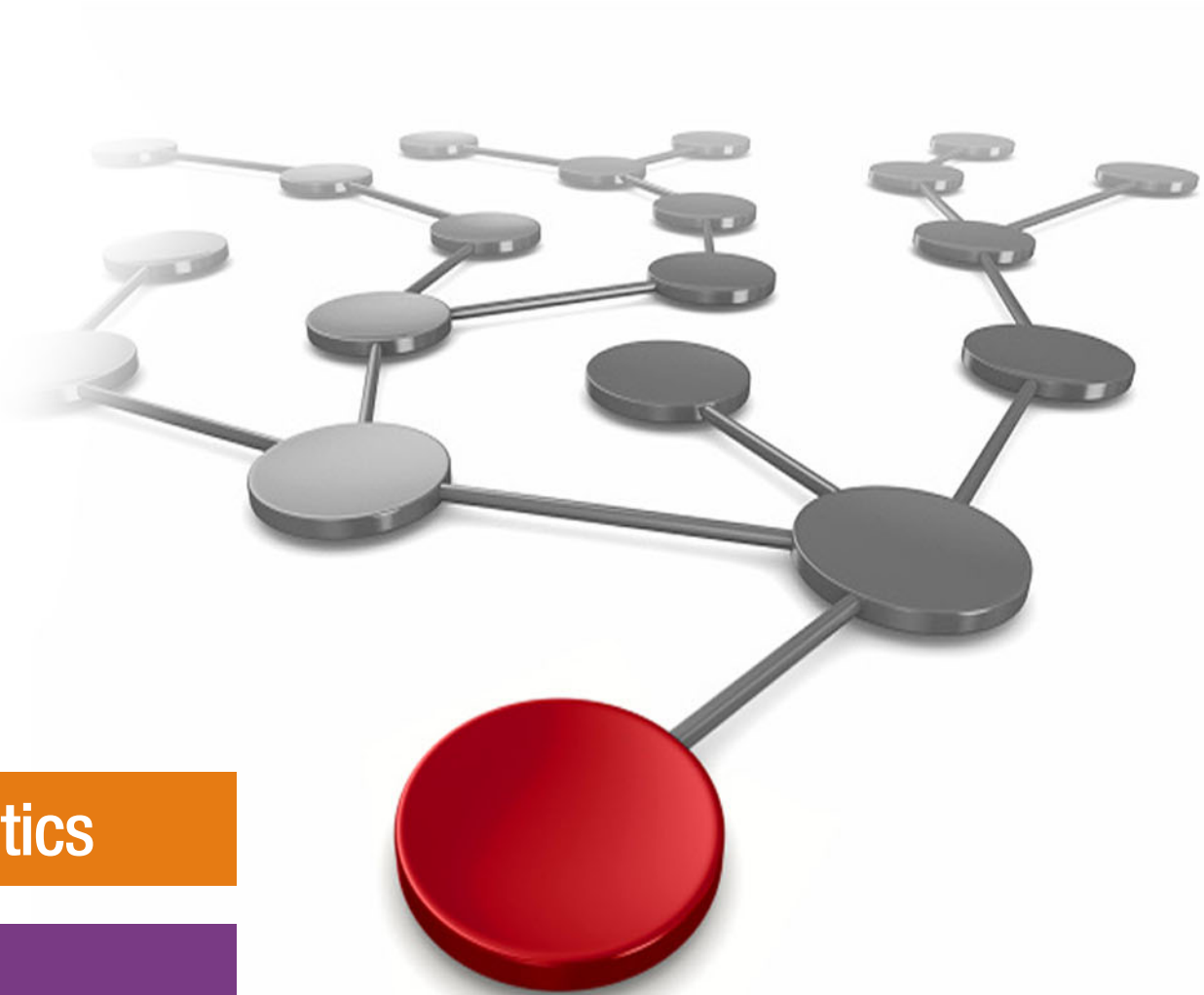


Big Data Analytics with IBM Cognos Dynamic Cubes

David Cushing

MaryAlice Campbell

Avery Hagleitner



 Analytics

Big Data

Big Data Analytics with IBM Cognos Dynamic Cubes

IBM® Cognos® Dynamic Cubes, which is a feature of the IBM Cognos Business Intelligence V10.2.2 software, complements the existing query engine. As explained in this IBM Redbooks® Solution Guide, it extends Cognos scalability to enable speed-of-thought analytics over terabytes of enterprise data, without being forced to rely on a new data-warehousing appliance. With this capability, which adds a level of query intelligence, you can unleash the power of your large enterprise data warehouse.

Figure 1 illustrates how the IBM Cognos Dynamic Cubes is integrated into the IBM Cognos Business Intelligence stack.

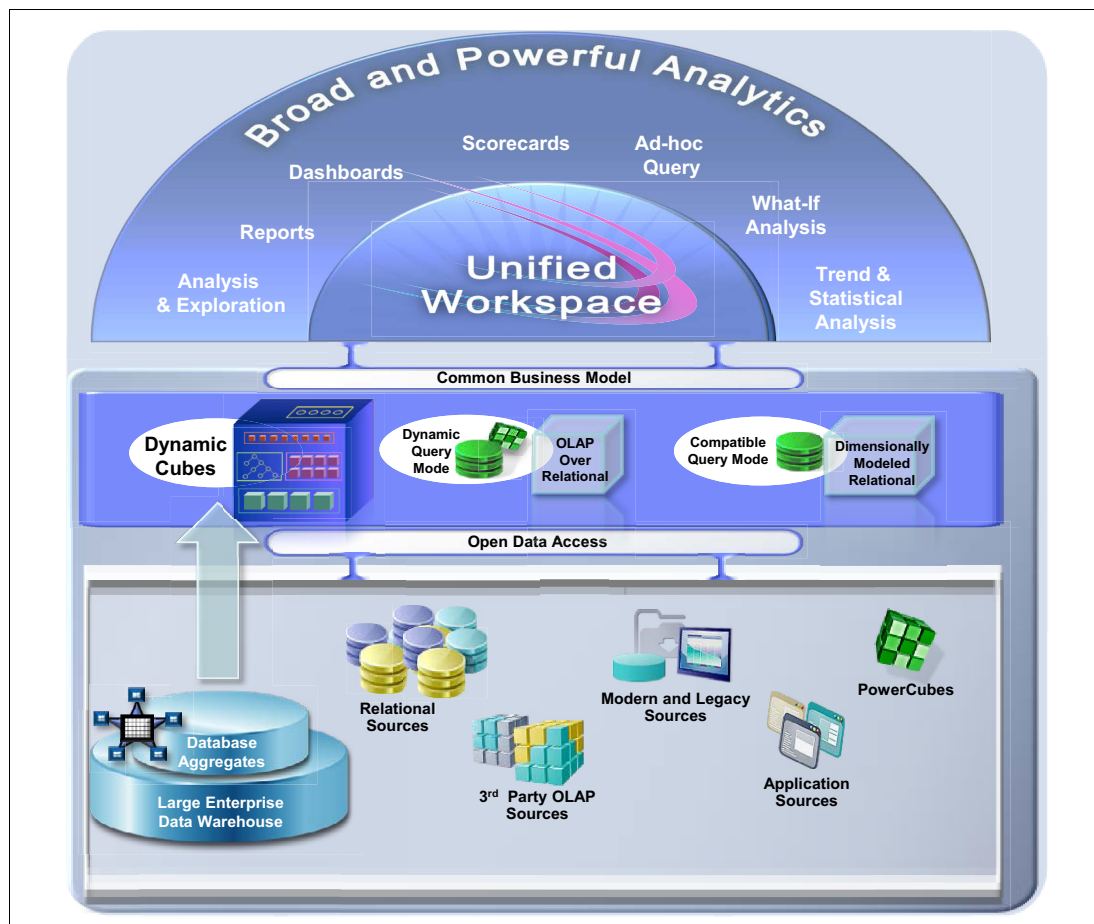


Figure 1 IBM Cognos Dynamic Cubes integrated into the IBM Cognos Business Intelligence stack

Did you know?

Every day, we create 2.5 quintillion bytes of data – so much that 90% of the data in the world today has been created in the last two years alone. Social media is now generating petabytes of data per day.

Business value

With social data generating petabytes per day, and instrumented devices becoming the norm, data volume growth is accelerating at an unprecedented pace. Big data is a growing business trend, with data from unconventional sources having the potential to be business disruptors. However, before the power of these new sources can be fully used, you must understand what is happening within your own business. Understanding your own business is added value of a data warehouse and is why taking full advantage of these data holdings is a critical first step to using these new sources of data. In addition, any organization that relies on instrumented infrastructures can maximize the efficiency of its operations. Analytics is key to accomplishing this type of optimization, leading to concrete business results.

Data warehouses are the recognized foundation for enterprise analytics. By using data warehouses, an organization can bring together cleansed data from separate sources of input, both internal and external, such as from partners or suppliers. Instead of garbage-in, garbage-out information to support decision-making, a consistent and consolidated enterprise-wide view of data from a business provides the foundation to improve your business. Building upon a trusted information platform for analytics is a key contributor to long-term business health. Not only do data warehouses enable higher quality information, they enable high-performance data access for analytic-style applications. IBM Cognos Dynamic Cubes technology helps in using the core strengths of an enterprise data warehouse and taking it to the next level of performance for analytics, making the deploying and tuning easier and faster.

Solution overview

The IBM Cognos Dynamic Cubes technology is meant to solve a specific but growing business problem, enabling high-performance interactive analysis over terabytes of data in an enterprise data warehouse. As data volumes grow, analyzing that data with speed-of-thought performance can be challenging. Even with modern data warehouse technology, some operations require significant computation or data movement. This computation or movement creates delays and reduces the satisfaction of business users who want to perform these analyses.

Various ways exist to accomplish performance over large volumes of data. From self-contained cubes to large in-memory appliances, different vendors are employing variations of similar methodologies to give business users timely response times. The Cognos Dynamic Cubes technology aims to give maximum flexibility in how memory is used to accelerate interactive analysis over terabytes of data so that you can evolve your deployments over time

Solution architecture

The IBM Cognos Dynamic Cubes technology is part of the IBM Cognos Business Intelligence query stack and is available with existing IBM Cognos entitlements. It provides a powerful means to enable high performance analytics over large data warehouses (see Figure 2).

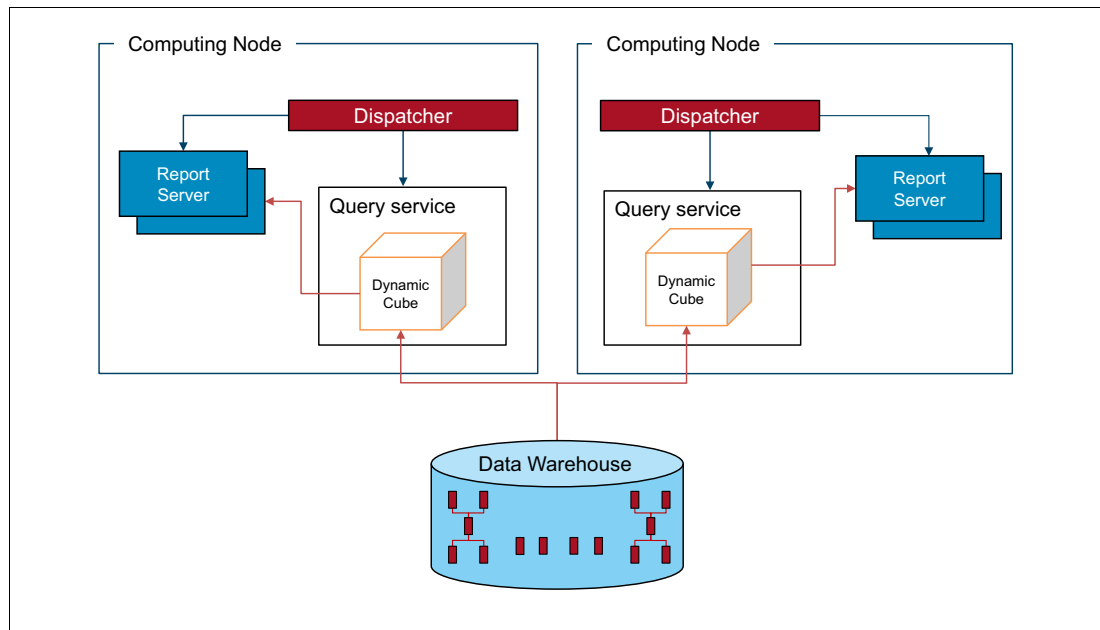


Figure 2 A high-level representation of the Cognos Dynamic Cubes architecture

The Cognos Dynamic Cubes solution consists of IBM Cognos Cube Designer (a modeling tool), a dynamic cube object in the administration environment (which becomes the data source), a package to enable Cognos BI client access to a dynamic cube, and the Aggregate Advisor (a wizard) that is started from within the Dynamic Query Analyzer.

IBM Cognos Cube Designer is a modeling tool that brings together the best modeling principles from past successful modeling technology, with a modern and extensible architecture. The first step to deploying Cognos Dynamic Cubes is to model with the Cognos Cube Designer.

After a dynamic cube is designed and deployed to the Cognos content store, it becomes available in the Cognos BI environment and is accessed through a corresponding package as an OLAP data source within any one of the Cognos BI client interfaces. A dynamic cube manages all aspects of data retrieval and leverages memory to maximize responsiveness, giving you full flexibility to manage what is in memory and when you want to refresh in-memory data. You manage dynamic cubes in the Cognos Administration Console.

A dynamic cube contains several in-memory elements to drive performance:

- ▶ Metadata members
- ▶ Aggregates
- ▶ Data
- ▶ Results sets
- ▶ Expressions

The Aggregate Advisor scans cube definitions and usage logs and then recommends both in-database and in-memory aggregates to improve performance. This approach helps to more easily address specific performance problems.

Usage scenarios

The Cognos Dynamic Cubes solution applies to the following usage scenarios.

Multigrain fact scenarios

A common requirement when creating a business application is to merge data from two or more areas of the business that have different scope or levels of granularity of the data. This might manifest itself as two or more star or snowflake schemas within a data warehouse that have different associated dimensions and, of those in common, different levels of granularity. For example, one fact table might include customer information and the other does not, and one fact table might contain data at the granularity of days while the other might contain data at the month level. This difference in fact grain might cause more difficulty in planning queries correctly if a report user included a level below the fact grain in the report.

With Dynamic Cubes, each fact table is modeled as a separate dynamic cube and these are in turn incorporated into a virtual cube. A virtual cube can be used to manage the presence of non-conformed dimensions and also common dimensions with different levels of granularity.

An example of such a use case is the need to compare actual versus plan sales data. The actual sales data might be at the granularity of day level whereas sales plan data might be recorded at the month level. By using a virtual cube that merges both of these cubes, you can make a query with day-level objects (or a member of that level, depending on the studio that you are using) against the sales facts. Then, you get the results and the expected null values for the planned sales facts. If you used a time dimension grain that was common to both fact tables, you get non-null values for measures from both fact tables.

In-database aggregates scenario

A second scenario addresses situations where a higher level of granularity of data is desirable for performance reasons. This is accomplished using a Dynamic Cubes feature called in-database aggregates. In-database aggregates define the measures, dimensions, and dimension grain by which queries can be routed to aggregate tables rather than to the detail fact table. Because in-database aggregate tables store fact data at a higher-than-detail level of granularity, the time necessary to aggregate values during the query can be lessened, improving performance. A query can be routed to the aggregate table if all the measures and dimension hierarchies of the query exist in the in-database aggregate definition. Not all of the dimensions and measures in the in-database aggregate must be in the query.

The objective of modeling an in-database aggregate is to establish rules by which a dynamic cube can detect when it can route a query to an aggregate table. This task is done by specifying a mapping from the identifiers in the dimensions and measures in the cube that have scope to the aggregate table, to the identifiers in the aggregate table. If necessary, mapping is done to its related tables in a rolled-up dimension schema.

This in-database aggregate routing directs a query only to the aggregate table for a query that uses objects from a dimension grain at or above the grain of the mapping between it and the aggregate table. Therefore, using objects from a grain below the mapping grain does not cause double-counting because, in this case, the query continues to route to the detail fact table.

Integration

IBM Cognos Dynamic Cubes is tightly integrated into the Cognos Business Intelligence stack, and its data can be surfaced through any of the Cognos client interface. As a result, existing customers can integrate Dynamic Cubes technology into their application environment without affecting existing users. Such users are already familiar with interfaces such as Report Studio, Cognos Workspace, and Cognos Workspace Advanced.

Different data requirements require different data solutions. One data path cannot be proficient at solving widely different data problems. Therefore, IBM Cognos has technologies that are built to suit specific application requirements. Table 1 can help you better understand the primary use case for each technology. However, carefully consider your individual application requirements when you make such a decision.

Table 1 Use cases for IBM Cognos data technologies

Cube technology	Primary use cases
IBM Cognos TM1®, in-memory cube technology with write-back support	<ul style="list-style-type: none"> ▶ It is optimal for write-back, what-if analysis, planning and budgeting, or other specialized applications. ▶ It can handle medium data volumes. The cube is run 100% in memory. ▶ Aggregation occurs on demand, which can affect performance with high data and high user volumes.
Dynamic Cubes, in-memory accelerator for dimensional analysis	<ul style="list-style-type: none"> ▶ It is optimal for read-only reporting and analytics over large data volumes. ▶ It provides extensive in-memory caching for performance, backed by aggregate awareness to use the power and scalability of a relational database. ▶ A star or snowflake schema is required in the underlying database (used to maximize performance).
PowerCubes, file-based cube with pre-aggregation	<ul style="list-style-type: none"> ▶ It is optimal to provide consistent interactive analysis experience to many users when the data source is an operational or transactional system, and a star or snowflake data structure cannot be achieved. ▶ The pre-aggregated cube architecture requires careful management, by using cube groups to achieve scalability. ▶ Data latency is inherent with pre-aggregated cube technology, where data movement into the cube is required.
OLAP Over Relational (OOR), dimensional view of a relational database	<ul style="list-style-type: none"> ▶ It is optimal to easily create a dimensional data exploration experience over low data volumes in an operational or transactional system, and where latency must be carefully managed. ▶ Caching on the Dynamic Query server helps performance. ▶ Processing that is associated with operational or transactional system affects performance.

Supported platforms

For information about software environments that are supported in IBM Cognos Business Intelligence V10.2.2, see the “Cognos Business Intelligence 10.2.2 Supported Software Environments” web page:

<http://www.ibm.com/support/docview.wss?uid=swg27042164>

Ordering information

Ordering information is shown in Table 2.

Table 2 Ordering part numbers and feature codes

Program number	Program name
5724-W12	IBM Cognos Business Intelligence V10.2.2
5724-W68	IBM Cognos Business Intelligence PowerPlay® V10.2.0
5724-W13	IBM Cognos Data Manager V10.2.0
5724-W20	IBM Cognos Mobile V10.1.0

Related information

For more information, see the following documents:

- ▶ *IBM Cognos Dynamic Cubes*, SG24-8064-01
- ▶ *IBM Cognos Business Intelligence V10.1 Handbook*, SG24-7912
- ▶ *IBM Cognos Dynamic Query*, SG24-8121

Also see the following web pages:

- ▶ Business Intelligence 10.2.2 documentation
<http://ibm.co/1zZb7hP>
- ▶ Dynamic Cubes Installation and Configuration Guide 10.2.2
<http://ibm.co/1FdEVI0>
- ▶ Dynamic Cubes User Guide 10.2.2
<http://ibm.co/1QIQVoG>
- ▶ Dynamic Query Analyzer Installation and Configuration Guide 10.2.2
<http://ibm.co/1b0VzBU>
- ▶ Dynamic Query Analyzer User Guide 10.2.2
<http://ibm.co/1H37JQm>
- ▶ Framework Manager User Guide 10.2.2
<http://ibm.co/1IAFpsD>

Authors

This Solution Guide was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center.

David Cushing is the Product Manager for IBM Cognos Dynamic Cubes, based in Ottawa, Canada. David has been with IBM Cognos for 25 years and joined IBM as a result of the acquisition of Cognos. David has a Master of Computer Science degree from Dalhousie University, Nova Scotia, Canada.

MaryAlice Campbell is a Senior Consultant and Business Analytics Technical Practice Leader at ISW, Australia. She has over 20 years of experience as a business analytics specialist. MaryAlice is an IBM Cognos BI veteran having gained experience with the early, pre-web versions of IBM Cognos PowerPlay and IBM Cognos Impromptu®; she contributed to beta and training programs, and worked with all subsequent releases. MaryAlice is also an IBM Certified Solution Developer, internationally recognized educator, and a Master Instructor of the IBM Analytics curriculum.

Avery Hagleitner is a Software Architect for IBM Cognos Dynamic Cubes in IBM Analytics Solutions at the IBM Silicon Valley Laboratory. Avery has over 14 years of software development experience at IBM. Her interests range from high-performance Java server applications to engaging graphical user interfaces. Her areas of expertise include business intelligence, data warehousing, and online analytical processing (OLAP). Avery holds a master's degree in Software Engineering from San Jose State University, California, USA, a Bachelor of Science degree in Computer Science and a minor in Psychology from the University of California, San Diego.

The project that produced this publication was managed by **Marcela Adan**, IBM Redbooks Project Leader - IBM International Technical Support Organization, Global Content Services.

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Stay connected to IBM Redbooks

- ▶ Find us on Facebook:
<http://www.facebook.com/IBMRedbooks>
- ▶ Follow us on Twitter:
<http://twitter.com/ibmredbooks>
- ▶ Look for us on LinkedIn:
<http://www.linkedin.com/groups?home=&gid=2130806>
- ▶ Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:
<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>
- ▶ Stay current on recent Redbooks publications with RSS Feeds:
<http://www.redbooks.ibm.com/rss.html>

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:


This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

This document, REDP-5265-00, was created or updated on September 11, 2015.

Trademarks

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. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US 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 <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

Cognos®	PowerPlay®	TM1®
IBM®	Redbooks®	
Impromptu®	Redbooks (logo)  ®	

The following terms are trademarks of other companies:

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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



REDP-5265-00

ISBN 0738454486

Printed in U.S.A.

Get connected

