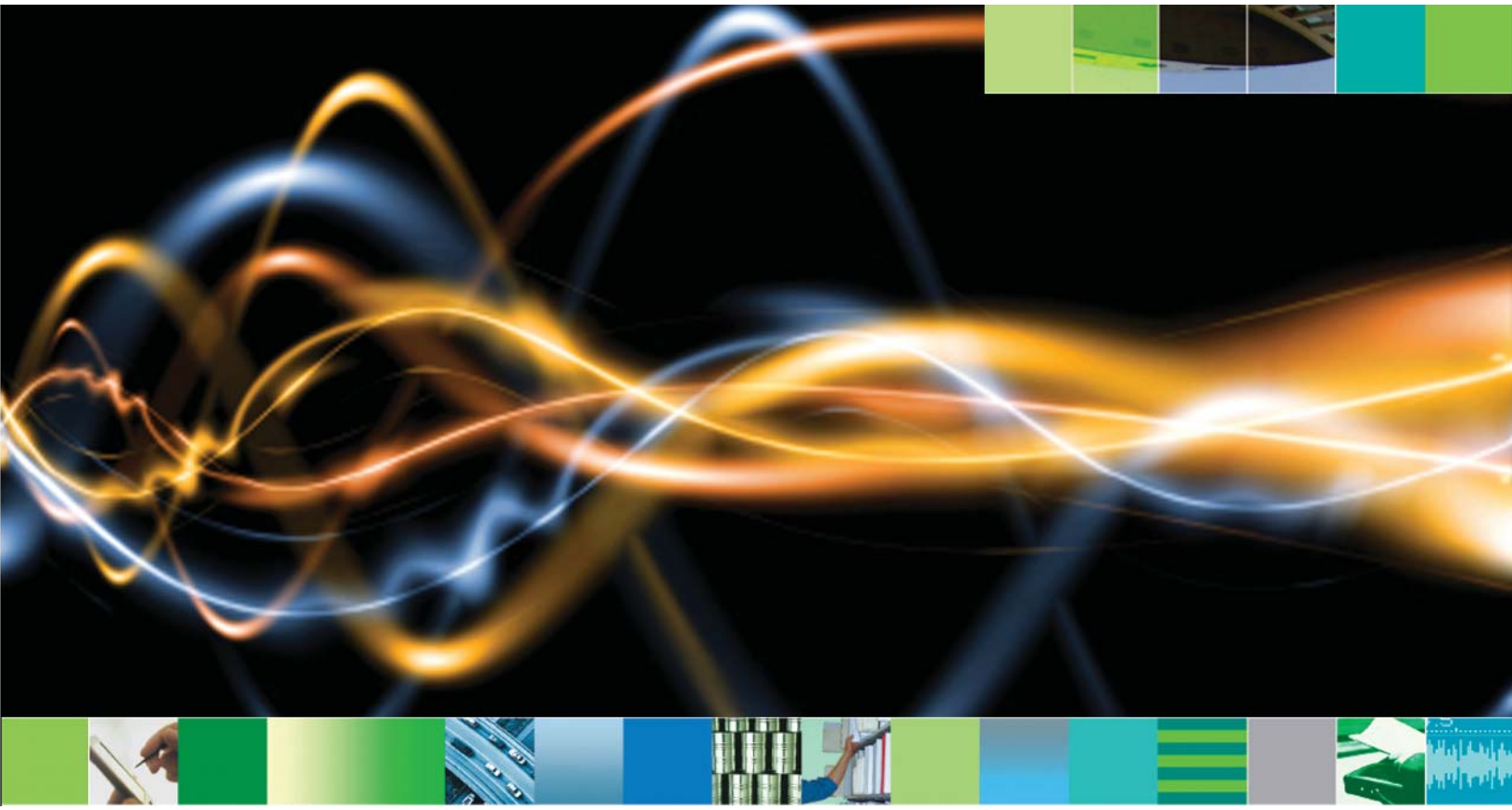
From the department to the enterprise
with IBM InfoSphere Warehouse
April 2009

Information Insight: From the Department to the Enterprise with IBM InfoSphere Warehouse



Information Management software

InfoSphere™ software

Table of Contents

2 *Limited reach, depth and responsiveness can hamper access to actionable insight*

5 *Achieve insight without boundaries with IBM InfoSphere Warehouse*

7 *Lowering the cost of ownership*

7 *Increasing the time to value*

8 *Greater storage savings with enhanced deep compression*

9 *Simplified analytics that enable business users to discover revenue opportunities faster*

9 *Increased system availability, reduced administration, and fewer unplanned outages*

10 *Operational insight with end to end support for XML documents*

11 *Simplify access to insight with embedded analytic services*

12 *Reach farther with unstructured analytics*

13 *Look deeper with integrated OLAP cubing services*

14 *Act faster with extreme workload management*

15 *IBM InfoSphere Warehouse in the real world*

16 *Enable IT managers to support the needs of the business through enhanced management capabilities*

17 *For more information*

Warehousing in the 21st century

As globalization gains momentum and the world of business becomes increasingly "flat," geographic and regulatory advantages are becoming less relevant. Thanks to similar technology and innovations, new ideas can be copied quickly—so companies cannot differentiate themselves with their products alone.

Infrastructure and personnel costs are spiraling, and ensuring compliance with strict regulations is costly but required. To stand out, companies must use the information collected throughout the enterprise to provide better service and more highly targeted product offerings. There is also a growing demand beyond the enterprise, specifically when the requirement is at the departmental level for a mission specific purpose, such as fraud and financial risk, customer loyalty or consumer advocacy. Often times these applications must reside on a separate information management system which more often than not is outside of the normal departmental budget.

Valuable information comes in many forms —from structured to unstructured, operational to transactional, real-time to historical— it is scattered throughout every enterprise. This information may reside in databases and data warehouses, e-mails and transaction logs, customer call logs, shopper behavior or repair orders or it may be XML data locked up inside transactional systems that cant be used or analyzed in your database. However, if you can unleash this data and leverage it properly, you will make better decisions and drive sales, improve processes and services, boost team productivity, reduce the risks inherent in doing business and streamline relationships with customers, trading partners and suppliers.

Highlights

Access to timely, accurate information is critical as enterprises strive to boost competitiveness and innovation.

To accomplish these objectives, business leaders must reach farther into critical business data and look more deeply to find hidden relationships, patterns and trends. They must also act more quickly using reliable, timely insight to drive sales, improve productivity and strengthen relationships with customers, partners and suppliers—all while keeping risks in check. However, barriers between information sources can make it difficult for people and applications to quickly and cost effectively use information to solve business challenges.

Limited reach, depth and responsiveness can hamper access to actionable insight

A variety of technological barriers may limit the types of information that businesses can leverage, the groups of people who can access the right insight and the speed at which it can be accessed. These barriers are the result of both historical IT system limitations and the way institutional processes are designed.

Because customer demands and competitive threats are evolving more rapidly than ever, status quo data warehousing solutions often cannot meet business requirements. These requirements may include giving employees insight from relevant real-time and historical information sources, using information to optimize strategic and tactical decision making or transforming key business information—and the way it is accessed, analyzed and acted upon—into competitive advantage.

To make better use of information at both the enterprise and departmental level users must overcome limitations within their existing environments. Current infrastructures typically have limited reach—which means that with the proliferation of transaction systems, business users must cast a wider

Highlights

Businesses have a significant advantage over competitors if they can quickly analyze all types of information and deliver actionable insights to executives and frontline decision makers.

net over the types of information they can combine and analyze to support better decision making. Unstructured information such as call center notes as well as structured information that resides in databases should be analyzed. Access to business intelligence (BI) tools and analytics should not be limited to high-level decision makers and specialized analysts. Instead, enterprises must deliver useful information to more people as part of everyday business processes.

In addition, existing information infrastructures often have limited depth. Business users must be able to get answers to more intuitive and complex ad hoc questions. IT teams require more flexible tools that can capture and deliver more types of information in the way that users need it—without driving up costs or creating management challenges.

Limited responsiveness also inhibits the effectiveness of many existing information infrastructures. Businesses have a significant advantage over competitors if they can quickly analyze all types of information and deliver actionable insights to executives and frontline decision makers. However, in older BI systems, only select users have access to limited types of data. Creating new information views requires technical expertise, which can restrict the speed at which tools can be modified to support business strategy. From a technical perspective, part of the challenge of supporting large numbers of users is enabling them to customize how they receive information based on their specific needs.

By overcoming these challenges and achieving access to insight on demand, businesses can achieve several critical objectives. They can microtarget small consumer segments and communicate with them about their individual needs and wants, as well as identify microtrends that can transform the enterprise. They may also be able to detect relatively small patterns of behavior that

Highlights

IBM InfoSphere Warehouse is designed to be the first comprehensive data server that enables businesses to centrally, accurately and securely analyze and deliver information as part of operational and strategic business applications.

Dynamic Warehousing from IBM is a new generation of business intelligence capabilities that enable organizations of any size to deliver business insights by integrating, transforming, mining and analyzing insights from both structured and unstructured information.

can have a significant influence on business. Most importantly, they can build competitive strategies around data-driven insights and generate impressive business results.

Achieve insight without boundaries with IBM InfoSphere Warehouse

Through dynamic warehousing, IBM can help companies extract insight from virtually any type of data—helping to deliver the right information at the right time so business leaders can make the right decisions. The IBM dynamic warehousing solution integrates data warehousing and business analytics to help define a company's central business concepts and the data required to support those concepts from an enterprise-wide perspective. It allows organizations to pull data from source systems that traditional business intelligence and data warehousing solutions cannot access, which makes it easier for IT organizations to support business requirements for actionable information—not just raw data, but data with intelligence behind it to help people take action and make sound business decisions.

IBM InfoSphere™ Warehouse is a complete, multipurpose environment that allows companies to access, analyze and act on operational and historical information—whether structured or unstructured—enabling them to get the insight and agility they need to consistently generate new opportunities, contain costs and satisfy customers. It is designed to be the first comprehensive data server that enables businesses to centrally, accurately and securely analyze and deliver information as part of operational and strategic business applications. Unlike traditional data warehouses and BI approaches that are complex, non integrated and rigid, InfoSphere Warehouse solutions simplify the processes of selecting, deploying and maintaining an affordable information management infrastructure that delivers the flexibility required by organizations to dynamically integrate and transform data into actionable business insight.

Highlights

InfoSphere Warehouse is available as a stand-alone software solution or as part of preconfigured, preintegrated, pretested InfoSphere Balanced Warehouse™ solutions.

From a technical perspective, the InfoSphere Warehouse platform provides a fully integrated environment built around IBM DB2® 9.7 server technology on Linux®, UNIX® and Microsoft® Windows® platforms. Common development and management user interfaces are designed to support application development, data modeling and mapping, SQL transformation, online application processing (OLAP) and data mining functionality. These features help to support the needs of business users to access, analyze and act on insights from virtually all types of information.

InfoSphere Warehouse is available as a stand-alone software solution or as part of preconfigured, preintegrated, pretested InfoSphere Balanced Warehouse™ solutions, which combine the simplicity of warehousing appliances with greater flexibility and reusability, better performance characteristics and greater functionality. Companies can choose from multiple versions—departmental through enterprise —that can help them to easily and flexibly grow the environment.

DB2 9.7 Enterprise Data Server, the core engine of InfoSphere Warehouse, can be deployed on servers of any size from small servers supporting small business or departmental applications through larger systems supporting 10's TO 100's of TB's data. It offers a highly scalable robust and resilient architecture that enables businesses to more effectively manage their information ensuring the required business functionality and service levels are met. Two of the major aims of this release are to provide businesses with a lower cost of ownership and faster time to value for their warehouse deployment.

Highlights

Lower Cost of Ownership

- *Deeper Compression*
- *Enhanced Workload Management*
- *Application Compatibility*
- *Simplified Warehouse Packaging*

Lowering the cost of ownership

DB2 9.7's enhanced data compression capability lowers the storage and processing capacity required for InfoSphere Warehouse users to manage information while enhanced workload management capabilities will increase the potential to consolidate a greater number of workloads and users onto the same platform.

DB2 9.7 support for XML data sources allows InfoSphere Warehouse to scale from natively analyzing gigabytes of XML documents through to 100's of TB's of XML documents by leveraging DB2's massive parallel processing engine. This combined with new compression support for XML documents drastically lowers the time to analysis and the cost of managing and analyzing XML information combining with the analysis of standard relational data using standard SQL query and reporting tools.

New cubing security and incremental data refresh capability maximizes the number of users and reports that can access a single cube. Thus reducing the number of cubes that needed to be managed and the amount of time required to manage the cubing environment.

Faster Time to Value Analytics

- *Accelerated Data Mining*
- *Enhanced OLAP*
- *Scalable Analysis of XML Data*

Accelerating the time to value

Database tooling and autonomic management enhancements improves the productivity of warehouse developers, reducing the time to value for warehouse implementations by allowing them to focus on key business deliverables with less time spent managing and setting up the warehouse platform. Additionally key enhancements support a broader set of database functionalities and language components to make it easier for analytical applications and new developers to migrate to the InfoSphere platform.

Highlights

Deep compression

- *Frees storage through a broader use of compression*
- *Reduce the footprint and cost of your Warehouse*
- *Workload management is now a standard feature of Departmental and Enterprise editions*
- *New Departmental Edition tailored to the needs of less complex organizational requirements.*

New data mining accelerators reduce the time it takes to model, test and deploy data mining underneath applications within the warehouse. These accelerators take users step by step through the process of modeling and visualizing their data through different mining algorithms.

New virtual cube capabilities will maximize the reuse of cubing models and the agility of developers and analysts to create new applications reducing the time it takes to deliver analytics to business users.

The following pages provide you with a more detailed view of the capabilities:

Greater storage savings with enhanced deep compression

InfoSphere Warehouse 9.7 includes new compression features that you can use to compress more types of data, further reduce your storage requirements, continue to improve I/O efficiency, and provide quick access to data from disk.

- XML data in the XML storage object of a table is now eligible for data row compression.
- Temporary tables are now automatically compressed with the DB2 optimizer to maximize storage savings and performance.
- Indexes can be compressed. If data row compression is enabled on a table, indexes on the compressed tables will be compressed by default.
- Data replication source tables can be compressed.

In addition, because some inline LOB files can now be stored in base table rows, you can use data row compression to compress them as well. When a LOB is smaller than a specified size, the LOB is now stored in the row of the base table instead of in the separate LOB storage object, increasing performance. Compression is included with InfoSphere Warehouse Enterprise Edition and is available through the Storage Optimization Feature license for InfoSphere Warehouse Enterprise Base Edition.

Highlights

Simplified analytics

- *Data mining wizards enabling a broader audience of users*
- *Data preparation and solution templates enable information as a service*
- *Multidimensional analysis with enhanced security features enabling direct analysis of DB2 data*

System Availability

- *Migrate data to new tables without loss of availability*
- *Improved online scalability through better node addition*
- *Better resiliency to errors enabling robust operational BI*
- *Easier workload management through a graphical user interface*
- *Automated threshold workload management that reconfigures for peak workloads*

Simplified analytics that enable business users to discover revenue opportunities faster

With InfoSphere Warehouse 9.7, users can now discover new relationships and patterns in information without being a data mining specialist. Improvements in data preparation and new data mining wizards and solution templates accelerate the creation of "ready-to-run" models.

- Support for time series analysis and prediction allow you to gain greater insight into transactional information, confidently forecast growth, and make better business decisions into the future.
- Support for virtual cubes and internal performance improvements in cubing services give business users more flexibility to perform multi-dimensional OLAP analysis in near real-time with large volumes of data.
- Fine grained security control in cubing services gives organizations the assurance to seamlessly deliver OLAP information across the enterprise wide.

Increased system availability, reduced administration, and fewer unplanned outages

Users can invoke a stored procedure to move the data in a table to a new table object of the same name (but with possibly different storage characteristics) while the data remains online and available for access.

Generate a new optimal compression dictionary when a table is moved. For system leveraging data partitioning a new node can be added online to minimize the impact on warehouse environments as the system grows. Database availability increased with enhanced resilience to errors. Remaining operational during peak business hours is the primary objective of every business. The enhanced resiliency of a DB2 instance, in the face of certain categories

Highlights

End to end support for XML documents

- *Directly store standard XML data into your data warehouse*
- *Directly query XML data in its native format*
- *Compress XML data to save on storage costs and improve performance*
- *Update XML data in its native format Improved performance when managing XML records*

of critical errors and traps, helps to achieve this objective. If a database instance needs to be recycled after the occurrence of a sustained critical error or trap, these short outages can now be scheduled after peak business hours.

Operational insight with end to end support for XML documents

For many organizations, the ability more effectively integrate XML into their information management environment has become a business necessity. With InfoSphere Warehouse organizations now have the ability to manage and analyze large volumes of XML data that were previously locked away in transactional systems.

InfoSphere Warehouse pureXML technology has been enhanced to support XML data in range partitioned tables, multidimensional clustering (MDC) tables, declared temporary tables, user defined functions, and partitioned database environments. To populate warehouse systems, the SQL Warehousing Tool now supports the extraction, transformation, and publishing of XML data. The performance of pureXML has also been improved. Multiple XML documents in a column can now be decomposed at the same time. In addition, write access is supported during creation or reorganization of indexes over XML data.

Highlights

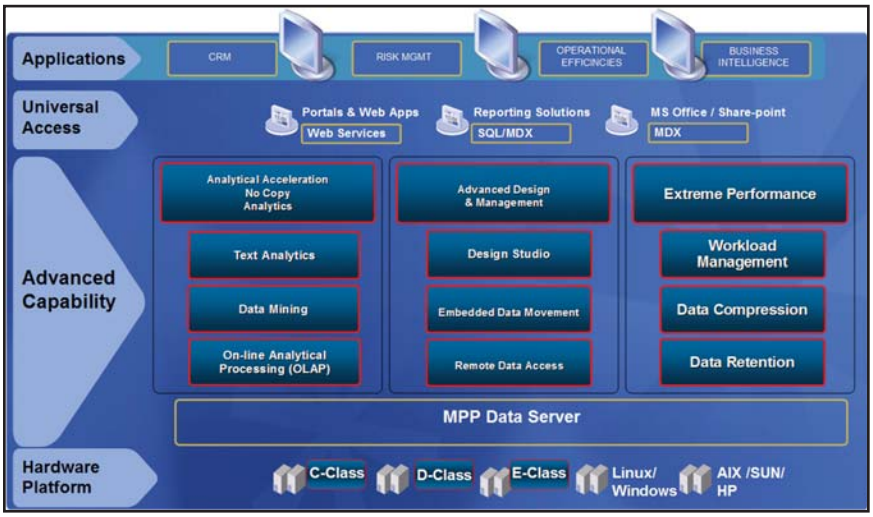
Figure 1: Unlike traditional data warehouses, IBM InfoSphere Warehouse embeds analytics functions within its architecture for enhanced insight, performance and simplified management.

Simplify access to insight with embedded analytic services

Traditional business intelligence tools are separate from the data warehouse, which creates two significant challenges:

- Data analysis requires that data be copied out of the warehouse, which introduces serious risks and limitations. The process may not capture all of the data or the most current data—and it requires copying the data again for every query, which is time consuming and does not allow for real-time analysis.
- Companies must maintain a set of tools that are separate from their data warehouse. They must upgrade or add new tools to the database and maintain the integration as the infrastructure evolves.

Figure 1 - Extensive InfoSphere Warehouse Possibilities



Highlights

By enabling businesses to extract knowledge and insights from previously untapped information, the unstructured analytics capabilities built into InfoSphere Warehouse can help to broaden the scope of information used in decision making.

The InfoSphere Warehouse eliminates these challenges by building BI tools directly into the platform (see Figure 1). Because this model eliminates the need to create and maintain duplicate copies of data along with the additional overhead necessary to maintain those copies, users are always working with the most complete and up-to-date information. Users can even access analytics tools from frequently used applications such as Web browsers or spreadsheets. This capability makes it easy to assemble applications, dashboards and other analytic views that deliver actionable information without requiring technical expertise.

The InfoSphere Warehouse also enables native processing in the warehouse and includes two basic annotators—dictionary and pattern-based extractors—out of the box. It is designed to simplify the process for assembling in-line analytics applications with point-and-click functionality. The platform also provides analytics application code that can be quickly modified and extended to address emerging business needs.

Reach farther with unstructured analytics

In the past, BI systems could only collect and analyze structured information. By enabling businesses to extract knowledge and insights from previously untapped information, the unstructured analytics capabilities built into InfoSphere Warehouse can help to broaden the scope of information used in decision making. Analysis of unstructured data such as call center notes can help organizations predict customer interest in associated or new products to generate cross-sell opportunities and growth.

InfoSphere Warehouse allows companies to extract interrelated knowledge from both structured and unstructured data by combining a set of search, text, analytics and visualization capabilities with unstructured analytics

Highlights

InfoSphere Warehouse offers outstanding scalability to handle large cubes of up to 11TB while also maintaining high performance and keeping costs down.

using existing tools and advanced search, text analytics and visualization capabilities. The platform automates key processes, making it easier to find issues in unstructured data and perform causal analysis relating them to associated structured data. Through this capability, companies can gain a better understanding of both immediate and long-term business issues and opportunities.

Unstructured analytics capabilities are pre-integrated and tested within InfoSphere Warehouse, which can help simplify implementation and maintenance and improve user adoption and return on investment (ROI). With a better understanding of the customer experience, companies can more effectively address issues encountered by their most profitable customers to reduce loyal-customer churn. In addition, these capabilities can help businesses avoid costly recalls and lawsuits with early warning reports.

Look deeper with integrated OLAP cubing services

Historically, access to business intelligence tools and reports has been limited to select "power" users and groups of decision makers. However, cubing services OLAP in DB2 9 allows companies to provide any employee with tools designed to help them analyze business data and generate business insights.

Cubing services OLAP allows users to ask intuitive and ad hoc questions that require multiple perspectives on data. This feature enables the exchange of information between systems, allowing business users to link multiple business variables together to perform deeper analyses than previously possible. In addition, rich presentation components allow visual analysis of data mining results, which can then be embedded into Web-based applications, customized and distributed to a broad range of users.

Highlights

Extreme workload management capabilities help to deliver real-time insight to large numbers of users without compromising performance.

InfoSphere Warehouse offers outstanding scalability to handle large cubes of up to 11TB while also maintaining high performance and keeping costs down. Through this feature, IT managers can maintain better control over unexpected resource consumption. The platform also uses standard interfaces to connect various analytic tools and support deep, multidimensional analysis of business information to generate insight while reducing costs. The platform supports mixed-mode environments with predictable results, helping to ensure that the database execution environment matches allocation of resources and active work to business priorities.

Act faster with extreme workload management

Data warehousing environments can be complicated, requiring specialized software and powerful hardware to deliver on service level agreements (SLAs). As the number of users who must access real-time business intelligence grows over time, enterprises often have difficulty balancing system performance, administrative workloads and costs. IBM helps companies address these challenges with InfoSphere Warehouse extreme workload management capabilities, which are designed to help companies consolidate the workloads from all business areas while better meeting business service level requirements for each group.

Ultimately, companies can save money as a result of providing users with a more holistic view of the company, a consistent view of information across functions and the ability to better analyze performance across divisions and activities. InfoSphere Warehouse includes robust monitoring and automation capabilities designed to help businesses proactively diagnose and resolve database issues to maintain optimum performance. In this manner, the platform can help companies apply business intelligence more broadly to improve operational business processes throughout the organization.

Highlights

Any mission, any size warehousing! The InfoSphere Warehouse delivers business value to customers throughout the world of all shapes and sizes.

IBM InfoSphere Warehouse in the real world

The New York Police Department's (NYPD's) innovative information management approach focuses on delivering relevant, comprehensive, insightful information to police officers from the moment they receive a new case. Today, NYPD uses InfoSphere Warehouse to help deliver crucial information to officers within moments of a reported crime—a brief window during which the right information can potentially lead to faster arrests. Officers now have fingertip, unprecedented access to information—such as 911 call logs, parole and probation files and recent crime reports—from scores of systems nationwide, plus the details on possible suspects in the vicinity. The ability to more easily explore all existing relevant information using analytics—including insight that is not captured in a typical data warehouse or transaction system—helps officers to act more quickly to save lives.

Highlights

Performance management features can help reduce support costs and maximize performance as warehouse deployments expand and new BI applications are added.

Enable IT managers to support the needs of the business through enhanced management capabilities

Complex environments such as data warehouses often mean complicated system administration, which reduces the time that IT managers can spend focusing on strategic projects to support business objectives. InfoSphere Warehouse helps organizations address this challenge by simplifying administration of data environments with an integrated software package that incorporates tools for deploying, managing and supporting data mining, modeling and analytics capabilities.

InfoSphere Warehouse is designed to allow IT employees to spend more time on activities that add value, rather than managing the data warehouse and BI systems. It can help streamline compliance through audit management features designed to simplify the process of determining who accessed or changed sensitive financial data and when. In addition, the platform can help organizations balance heavy data loads and query times by providing a quick, security-rich means for archiving historical data on an archived database.

Through new InfoSphere Warehouse Performance Suite features such as workload management, IT managers can better understand the effects of different applications and users on resources and performance to enable better system configuration and optimization. In addition, performance management features can help reduce support costs and maximize performance as warehouse deployments expand and new BI applications are added.

Graphical tools for complex transformations, native connectivity to non-DB2 sources, advanced workflow control and scheduling and automation of data mining can help simplify warehouse integration. In addition, unique storage

Highlights

Take a tour of the InfoSphere Warehouse Value Matrix to learn about the multiple levels of value that the IBM InfoSphere Warehouse can deliver to your company.

ibm.com/software/data/infosphere/warehouse/values/

optimization technology can dramatically reduce the space associated with storing relational data—helping to reduce costs and improve disk utilization and query speed.

For more information

To learn more about IBM InfoSphere Warehouse, please contact your IBM representative or IBM Business Partner or visit ibm.com/software/data/infosphere/warehouse

© Copyright IBM Corporation 2009

IBM Corporation
Software Group
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
April 2009
All Rights Reserved

IBM, the IBM logo, ibm.com, InfoSphere, DB2, Cognos, and pureXML are trademarks or registered trademarks of International Business Machines Corporation in the United States, other coun-tries, 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 trade-marks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at ibm.com/legal/copytrade.shtml

Intel and Itanium are trademarks or registered trade-marks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Linux is a registered 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.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates. The information contained in this documentation is pro-vided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this documentation, it is provided “as is” without warranty of any kind, express or implied. In addition, this information is based on IBM’s current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this documentation or any other documentation. Nothing contained in this documentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM (or its suppliers or licensors), or altering the terms and conditions of the applicable license agreement governing the use of IBM software.

Each IBM customer is responsible for ensuring its own compliance with legal requirements. IBM customers are responsible for ensuring their own compliance with legal requirements. It is the customer’s sole responsibility to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer’s business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law.

♻ Printed in the United States of America on recycled paper containing 10% recovered post-consumer fiber.

IMW10925-USEN-02