Information integration software solutions
White paper

DB2 Information Management Software

Creating a flexible infrastructure for integrating information.

by Holly A. Hayes
IBM Software

October 2003

Contents

Executive summary

Integrating information across and beyond the enterprise is a competitive mandate. Initiatives such as customer relationship management, supply chain management and business intelligence are based on successfully integrating information from multiple data sources, both structured and unstructured. Because of organizational, operational or cost constraints, these data sources do not generally lend themselves to being fully replicated or consolidated within a single database. Accordingly, there is an increased demand for federated access to distributed sources. As enterprises evolve into on demand businesses, it becomes increasingly important to integrate diverse data flexibly and efficiently. Information integration is fundamental for companies to reap the benefits of e-business on demand.

Information integration technology enables integrated, real-time access to traditional and emerging data sources, transforms information to meet the needs of business analysts and manages data placement for performance, currency and availability. As a result, organizations can increase efficiencies, better serve customers and suppliers, make more informed decisions and respond more quickly to new opportunities or threats.

This white paper is intended for IT decision makers and influencers who are evaluating integration technologies. The paper describes the various challenges that businesses face today, the on demand operating environment and the role of information integration. In addition, this paper profiles ways in which the DB2® family of information management products from IBM addresses information integration needs.

Integration challenges

Integrating information for an enterprise is a daunting job. IT executives must deal with pressure from the top to deliver results quickly, meet end-user expectations for ease of access to highly diverse data sources, plan around technology limitations and adapt to a constantly changing environment.

In today's competitive environment, businesses want to:

- Reach a broader customer base over the Internet while integrating new Web applications with existing core business processes.
- Enhance the value of portals with information from critical e-business systems, improving the productivity of business users.
- Speed product delivery by integrating order processing with internal application areas, such as manufacturing and shipping, as well as with suppliers and other trading partners.
- Capitalize on emerging opportunities more quickly by correlating information across competitive analysis, analyst research, sales information and customer demographics.
- Integrate all customer information across the enterprise with purchased demographics data to personalize interactions for improved customer loyalty and revenue generation.
- Streamline information flow between people, processes and applications to minimize unnecessary work and delay.

An Internet generation

The Internet has revolutionized customer expectations for service and information access. Through search engines such as Google, Lycos and Yahoo!, people have access to a collection of information on virtually any topic. Response times are typically dictated more by connection speed than by any other factor.

Businesses have attracted customer attention and loyalty by bringing together complete views of client portfolios, adding value by incorporating interesting and relevant content. The bottom line: Customers expect fast, online access to a holistic view of their portfolios with broad-based, value-added content.

Information explosion and diversity

There is no shortage of content. Digitized information is growing rapidly, seemingly beyond the ability of businesses to manage and leverage it. Industry analysts at the School of Information Management and Systems at the University of California, Berkeley, predict that more data will be generated from 2001 through 2003 than in all of recorded history. According to the Berkeley researchers, "The world produces between 1 and 2 exabytes [1 and 2 billion gigabytes] of unique information per year, which is roughly 250 megabytes for every man, woman and child on earth."[1]

The growth in information, combined with the diversity of information sources, further complicates the retrieval of useful information. Businesses must access not only traditional application sources such as relational databases, but also Extensible Markup Language (XML) documents, text documents, scanned images, video clips, news feeds, Web content, e-mail, analytical cubes and special-purpose stores, both internal and external. Because of organizational or operational constraints, information from diverse and distributed data sources does not generally lend itself to being fully replicated or consolidated into a single database. Yet hidden information can be uncovered, opportunities more readily recognized and customers better served when information is correlated. Giga Information Group estimates that at least 30 percent of all new e-business applications face the problem of integrating multiple data sources.[2]

[1] P. Lyman, H. Varian, J. Dunn, A. Strygin, K. Swearingen, "How Much Information?", University of California, Berkeley, October 2000, *http://sims.berkeley.edu/research/projects/how-much-info/*

[2] Giga Information Group, Emerging Internet Data Integration Solutions, November 2000.

Providing coherent access to diverse data is a significant obstacle for most businesses. Not only are technology solutions limited, but there also is the potentially larger problem of gaining corporate agreement on a common taxonomy.

Technological challenges to integrating information

Technology has struggled to keep pace with integration requirements. Businesses cobble together integration solutions to meet urgent needs, only to realize later that the solution lacks scalability, availability and flexibility.

Technology vendors in a myriad of markets—such as enterprise application integration, data warehousing, enterprise content management, portals and application servers—have begun to shift their focus toward the overall integration problem. This makes it more difficult for a customer to choose the best technology to meet the business needs. Moreover, the niche orientation of point products often makes it difficult to leverage the benefits of one implementation in a subsequent project. Customers may find themselves integrating the integration solutions.

A changing environment

The integration job is never finished. IT environments are in a constant state of flux. New applications come online. Release-level changes to packaged applications can cause a ripple effect throughout the infrastructure. There is always the next new tool or technology to try. Investments must be made with an eye to the future. Thus, organizations are emerging within businesses to focus on an integration architecture. Whether called Information Management, Integration Services or Data Architecture, specialized departments within companies are tackling the issue of integrating the business, and defining an integration architecture and infrastructure that will provide the foundation for their future.

The on demand era

A new era is emerging. Businesses are trying to become more *on demand*. An on demand business is an enterprise whose business processes—integrated end-to-end across the company and with key partners, suppliers and customers—can respond with speed to any customer demand, market opportunity or external threat. To achieve that vision, an enterprise will need an IT infrastructure whose priorities are aligned with business goals and which enables the business to be more flexible and responsive. We call this infrastructure the on demand operating environment. It is designated to help enterprises reduce costs, improve asset utilization and address new business opportunities.

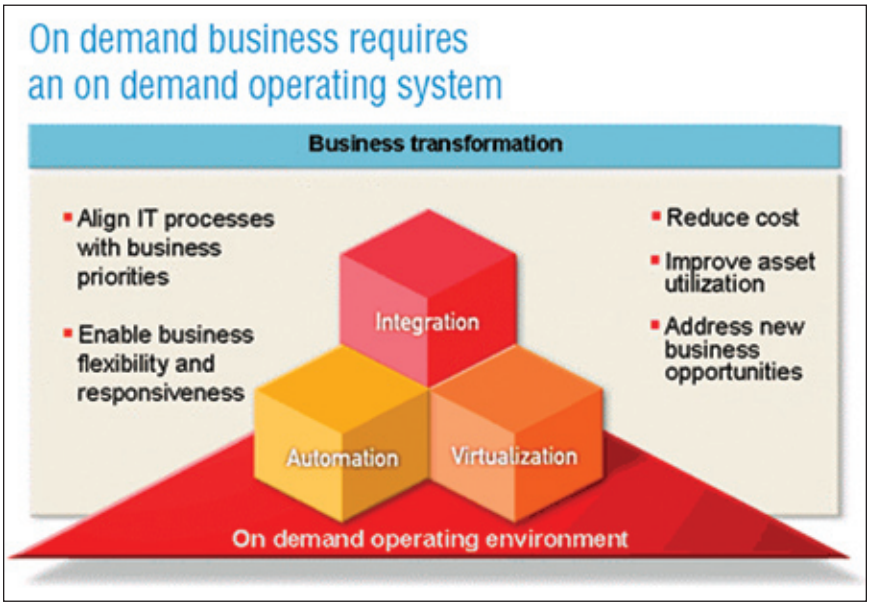

Figure 1. The on demand operating environment helps businesses align IT processes with business priorities to enable business flexibility and responsiveness.

The on demand operating environment is an integrated platform, based on open standards, that enables rapid deployment and integration of business applications and processes. It is:

- Integrated—enabling the efficient and flexible combination of resources to optimize operations across and beyond the enterprise.
- Automated—reducing systems management complexity to enable better use of assets, improve availability and resiliency, and reduce costs based on business policies and objectives.
- Virtualized—providing a single, consolidated view of and easy access to all available resources in a network, no matter where they reside.

Integration revolves around people, processes and information. Different integration capabilities are necessary for different classes of integration problems. For example, an insurance company might use information integration to access all claims information related to a specific customer (such as policies held, payment status, previous claims, fraud information, and so on); a portal to deliver the processes and information required for claims processing to the claims adjuster (customer policy details, first notice of loss details, comparable incidents, etc.); and process integration to streamline and automate the claims process from first notice of loss through to settlement. As in this example, the best solution often utilizes several capabilities, which emphasizes the need for switching easily from one to another.

While competitors may provide only niche integration, IBM can deliver a comprehensive integration platform with offerings that work together seamlessly. IBM has more than 30 years' experience in building and evolving the base technologies for middleware and enabling these technologies to work together in thousands of different business environments.

A single view of data

IBM envisions an information integration infrastructure that provides the application layer with a unified view of the data it must access, regardless of differences in data format, data location and access interfaces. The evolution of data management software is not simply about managing single-instance data stores, but about providing value-added integration across all forms of data, dynamically managing data placement to match availability, currency and performance requirements, and providing autonomic features that continue to reduce the burden on IT staffs for managing complex data architectures. Ultimately, it is about getting the maximum value from your information assets. Based on ongoing research and proven data management technologies in areas such as relational data, XML, content management, federation, search and replication, IBM is developing the infrastructure shown in Figure 2.

Figure 2. The IBM infrastructure provides flexible access based on a range of programming models, a rich set of integration features and interoperability with IBM's overall business integration framework.

Flexible access

IBM's vision is to provide flexible access to an information integration infrastructure through industry-standard interfaces. Client access may be through ODBC, JBDC, Web services, native client or asynchronous client interfaces. Supported query languages will include:

- Structured query language (SQL), the industry's most mature and powerful query language with widespread adoption in the marketplace.
- XQuery, the emerging standard for access to XML data, currently being standardized through the World Wide Web Consortium (W3C).
- IBM DB2 Content Manager object-oriented application programming interfaces that support the content management lifecycle, including rich text and image query.

Regardless of the client access and query language used, the application should be able to access all of the data connected through the integration server. This combination of client access and query language flexibility makes it possible for existing development and analytical tools to take immediate advantage of the broader data access and integration features provided by the integration server. It also allows the infrastructure to plug into service-oriented architectures using Web services; to provide asynchronous clients for easy integration with workflows or scheduling long-running queries; and to extend your investment in current and new application infrastructures.

Rich features

The IBM information integration infrastructure will enable real-time integration of diverse and distributed data as if it were a single source, no matter where it resides. The key features of the infrastructure include the ability to federate, search, cache, transform and replicate disparate data:

Federate. IBM provides industry-leading federation over diverse data sources. Federation refers to the ability to view and manipulate a collection of data sources as if they were a single source, while retaining their autonomy and integrity. The resources may be uniform or diverse, collocated or distributed, depending on the implementation. IBM's federation engine provides:

- *Transparency, which helps mask from the user the differences, idiosyncrasies and implementations of the underlying data sources, to make the set of federated sources appear as a single system.*
- *Heterogeneity, which implies the ability to federate highly diverse types of data, including structured data (e.g. relational databases), semi-structured data (e.g. XML documents) and unstructured data (e.g. free-form text).*
- *Extensibility, such that the federation can be extended to almost any data source. Specifically, the extensibility has been designed to minimize the effort required to integrate a new source, yet offer the flexibility to provide the necessary information to optimize query access.*
- *Rich functionality, which includes the functions available through the supported query languages, compensation for missing functions in the back-end data sources, plus the ability to include source-specific capabilities into the query language seamlessly.*

- *Autonomy for data sources such that they can be federated with little or no impact on existing applications or systems.*

- *Performance characteristics that make federated query a real-world option. IBM, with more than 25 years of research and development, as well as patented optimization technology, can make federated query possible.*

Search. IBM's infrastructure will provide advanced search and query capabilities including the ability to crawl the Web, index documents, federate search results from multiple search engines, categorize and summarize text documents for intelligent access, and understand semantics. In 2002, IBM formed the IBM Search and Text Analysis Institute to unify and accelerate IBM's research and deployment of advanced search and mining functions through an integrated architecture. The results will fuel IBM's information integration platform as well as other IBM offerings.

Cache. A range of caching strategies is required to provide adequate performance, currency and availability characteristics to requesting applications. IBM's information integration infrastructure will support placing and managing data at multiple points in the data hierarchy to improve performance. Beyond simple caching, this is policy-based data placement and management.

Transform. The infrastructure must provide rich transformation features to facilitate analysis, interchange or presentation.

Replicate. Replication is required as a fundamental characteristic of an information integration infrastructure. It complements the distributed access features, enabling management of centralized data stores, and provides the necessary infrastructure for managing data caches efficiently.

Comprehensive business integration

As noted earlier, information integration is just part of an overall business integration infrastructure. To support a business's evolving requirements, this infrastructure is complemented by robust data and content stores, as well as additional integration technologies, and is based on industry standards.

Storing the data: IBM today offers leading relational database management systems and content management systems, and is leading the industry in delivering integrated XML support. Beyond current capabilities, which are based on the relational data model, the XML store must completely embody and exploit the XML data model. An XML registry will provide easy management of the vast number of XML artifacts, such as XML schema documents, document type definitions (DTDs) and Web services description documents that can be generated as the quantity and variety of XML data increases.

Leveraging complementary integration technologies: Key to developing a corporate integration infrastructure is the ability to easily leverage appropriate integration technologies, together or separately. IBM continues to focus on integration across the IBM Software Group portfolio—and most importantly with the WebSphere® portfolio—to deliver a comprehensive business integration infrastructure.

Support for industry standards: IBM continues to be at the forefront of industry standards development and adoption, facilitating broad interoperability among vendor tools.

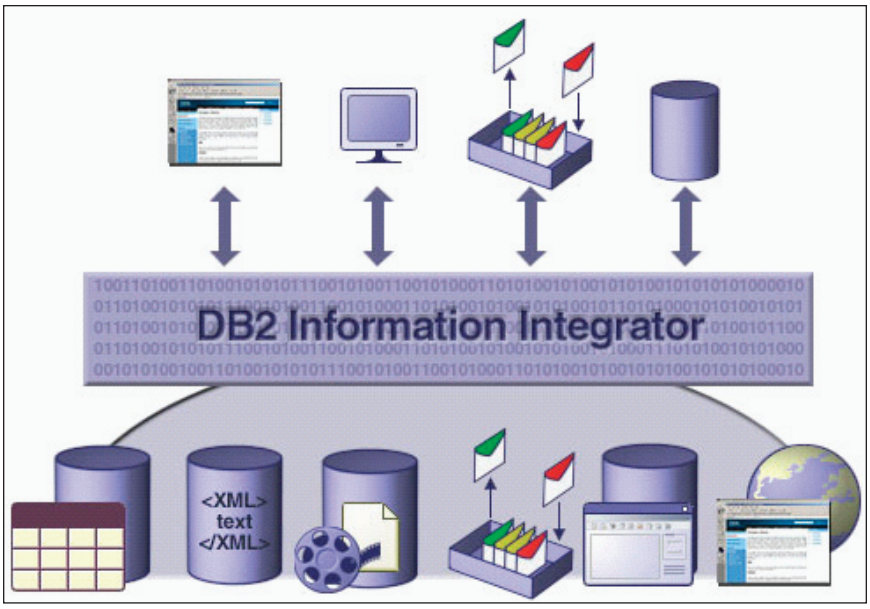

Figure 3. DB2 Information Integrator products deliver integrated access to diverse, distributed and real-time data as if it came from a single data source.

Introducing IBM DB2 Information Integrator

IBM DB2 Information Integrator software provides the foundation for a strategic information integration framework, as shown in Figure 3. Such a framework helps customers to access, manipulate and integrate diverse and distributed data in real time. The portfolio consists of:

- IBM DB2 Information Integrator, V8.1, a new product based on DB2 information management technology.
- IBM DB2 Information Integrator for Content, V8.2, formerly IBM Enterprise Information Portal.

Each of these products enables customers to abstract a common data model across diverse and distributed data and content sources, and to access and manipulate them as though they were a single source. Each supports a user community defined primarily by the data its members access and the development community they support. This product set supports the predominantly read-access scenarios common to enterprisewide reporting, knowledge management, business intelligence, portal infrastructures and customer relationship management.

With the DB2 Information Integrator family of products you can:

- *Choose the data access strategy to match the business value.* Consolidating data and making it local to the application makes application development easier and provides better data access performance and availability. However, it also introduces the burden and cost of moving data from place to place, storing the data and managing its synchronization. DB2 Information Integrator V8.1 delivers centralized access by providing replication and caching to support performance and availability requirements. Alternatively, it may be better to access data in place when there is wide diversity in the data accessed, it is impractical or too expensive to replicate the data, or when the data is owned outside the enterprise.

- *Integrate data and content without moving the data or changing the platform.* The DB2 Information Integrator family lets you access diverse and distributed data no matter where it resides, as though it were sourced from a single location. It provides a broad range of data source access out-of-the-box, covering structured and unstructured data across and beyond the enterprise. With more information available more easily, you have the opportunity to gain greater return on existing information assets.

- *Make more progress, more quickly and at a lower cost.* With DB2 Information Integrator products you can more easily and quickly develop a new generation of composite applications that require efficient integration of disparate data. Developers can choose whether to use an SQL or DB2 Content Manager programming model based on the product selection. And now, you have a practical way to integrate diverse relational data and combine it with unstructured data from content repositories, the Web, spreadsheets or other sources. Thus, the DB2 Information Integrator offerings can help you speed project deployment, leverage existing skills over a broader range of projects and reduce ongoing maintenance costs.

DB2 Information Integrator: a server for federated data and replication

DB2 Information Integrator, V8.1, targets at the application development community familiar with relational database application development. Applications that use SQL, or tools that generate SQL (such as integrated development environments and reporting and analytical tools), can now access and manipulate distributed and diverse data through a federated data server. This product is most appropriate for projects whose primary data sources are relational data augmented by other XML, Web or content sources.

DB2 Information Integrator, V8.1, is based on the DB2 technology infrastructure, leveraging IBM's prior investments in products such as IBM DB2 DataJoiner®, IBM DB2 Relational Connect and IBM DiscoveryLink® (as well as subsequent product enhancements). DB2 Universal Database™ consists of a modern database architecture known worldwide for its scalability and extensibility.

DB2 Information Integrator includes the ability to federate, search, cache, transform and replicate data. As a federated data server it provides out-of-the box access to DB2 Universal Database, DB2 Informix® Edition products, as well as databases from Microsoft, Oracle, Sybase and Teradata. In addition, it can access semi-structured data from WebSphere MQ messages, XML documents, Web services, Microsoft Excel, flat files, ODBC or OLE DB sources, plus a variety of formats unique to the life sciences industry. Integrated support for IBM Lotus® Extended Search provides the solution's broad content access to a variety of content repositories, including DB2 Content Manager, as well as e-mail databases, document repositories, third-party Internet search engines and LDAP directories. A developer's kit extends the federation capability to virtually any data source.

The solution provides search and query access through a standard SQL API, and combines the broad content access of Lotus Extended Search with the precision of a relational engine. Two search approaches for text are available:

- The ability to create a global index for back-end relational stores. Using this approach, text search semantics—such as fuzzy search, thesaurus support and search within sections—may be used within the query.

- A brokered search architecture that does not require the creation or maintenance of a central index to access content across multiple sources. The extended search engine translates each full text query into the native query language of the target data source.

The query may produce standard SQL answer sets or XML documents. XML documents can additionally be validated and published to a WebSphere MQ message queue.

The DB2 Information Integrator optimizer has been extended significantly to support distributed federated query processing.

- Query rewrite, a powerful phase of query optimization that transforms poorly written input queries into semantically equivalent forms for improved performance, is aware of underlying data sources, and will restrict or enable particular transformations based on their applicability to the particular data source.

- Pushdown analysis is a newly introduced phase of query processing that decides how much of a particular query can be evaluated by each particular back-end server, and how much compensatory processing needs to occur on the DB2 Information Integrator system.

- Cost-based optimization creates a query execution plan based on cost estimates, which now include standard statistics from source data (for example, cardinality or indexes), data server capability (such as join features or built-in functions), data server capacity, I/O capacity and network capacity.

- Statement generation, which generates an executable plan based on the results of the cost-based optimizer, has been extended to produce efficient DBMS-specific SQL for SQL-speaking sources.

- The query run-time engine has been extended to drive execution of a query over both local and distributed information, allowing for functional compensation and providing a coherent virtual database view.

- The first release of federated caching provides administrator-managed caching of integrated views across relational database back ends. The optimizer automatically routes queries to the cache to satisfy queries when appropriate.

A rich set of transformation features includes standard SQL functions, such as string manipulation, arithmetic calculations, statistical computations, online analytical processing functions and procedural logic. Type-specific features—such as the application of scoring algorithms or chemical similarity searches—further enhance this already rich set of transformations. Extensible stylesheet language (XSL) translations facilitate document interchange and dynamic style matching to diverse display characteristics. User-defined functions enable customers to standardize virtually any function for any data type. In addition, the ability to access Web services as a built-in function means that any Web service, such as a currency conversion, can become an embedded transformation function.

DB2 Information Integrator, V8.1, also includes a replication server for mixed relational databases. Customers can replicate data between IBM (DB2 Universal Database and DB2 Informix Dynamic Server), Microsoft, Oracle and Sybase. DB2 Informix XPS and Teradata can also be replication targets. Customers can configure a variety of topologies, latencies and consistency characteristics.

In a recent CRN article, "IBM: Xperanto Rollout to Start In Early 2003",[3] industry analysts lauded IBM's upcoming offering:

- IBM's big advantage will be its reliance on its own optimization logic "which is second to none," and its ability to handle structured and unstructured data. [Doug Laney, Vice President Application Delivery Strategies at Meta Group]

- "IBM will raise the bar on enterprise information integration. It has so much more to offer than small companies, and it has broadened the category to include unstructured data," says Philip Russom, research director of Data Integration at Giga Information Group.

[3] For more information, please visit: *http://www.crn.com/sections/BreakingNews/breakingnews.asp?ArticleID=39187*

DB2 Information Integrator for Content: federated access to diverse content

DB2 Information Integrator for Content, V8.2, targets the content application developer who needs to search for and access text and non-text information across a wide range of content sources. Providing seamless reach into diverse data environments, DB2 Information Integrator for Content represents a renaming and repositioning of the IBM Enterprise Information Portal offering.

DB2 Information Integrator for Content offers a rich set of integration features, such as connectors to diverse content sources, sophisticated information mining and advanced workflow. To speed implementation of content integration projects, DB2 Information Integrator for Content provides access to a variety of data sources out of the box—all of which can be federated in a single search. The connectors include access to the DB2 Content Manager family and other content repositories, Lotus databases, relational databases and wide-ranging content available with Lotus Extended Search.

Additionally, DB2 Information Integrator for Content includes a sophisticated information mining capability that uses Web crawling and text-mining algorithms to provide structure to unstructured content. The mining algorithms include the ability to identify the language in which a document is written, identify features within documents such as names, classify documents according to a defined taxonomy, group documents by category and summarize documents. By building their knowledge about enterprisewide information, businesses can reap additional returns from existing content assets.

Finally, DB2 Information Integrator for Content provides an advanced workflow application to enable businesses to increase productivity, reduce production times, and improve communication and collaboration. Using a graphical workflow builder, developers easily define workflow processes across the enterprise.

The DB2 Information Integrator portfolio addresses the need to access and integrate structured and unstructured data. Supporting today's popular programming models, each offering enables businesses to quickly capitalize on their existing skills and tools infrastructure in enterprise content management and SQL-related deployments. IBM has a compelling vision, a significant research and development investment, and a roadmap for the product family that represents the evolution of information management technology.

Completing the solution

DB2 information management software represents a rich set of offerings that deliver a complete solution for information management. These proven solutions include DB2 Universal Database, DB2 DataPropagator™, DB2 Content Manager and DB2 Warehouse Manager.

DB2 Universal Database provides the foundation engine for storing, managing and federating data. DB2 includes a rich set of statistical and analytical functions and offers the best optimization technology in the industry, including distributed optimization and automated caching functions. The DB2 Extenders™—including DB2 XML, Text, Net.Search, Audio, Visual, Image and Spatial Extenders—provide data type-specific extensions to query, access, update and manage various data objects. For example, with DB2 Extenders, you can manipulate XML documents, query by image shape or color, or query based on proximity to a given location. Consider Satellite Records, which plans to use IBM DB2 XML Extender to process XML-based documents from distributors and store them in DB2. The company will thus be able to easily move vinyl record information it receives from distributors directly to its own Web site. Steve Shapero, IT director at Satellite, estimates that DB2 XML Extender will ultimately save the company 40 percent in development time and costs by providing a framework for its extranet, reducing the amount of new code that the company would otherwise need to write.[4]

[4] Satellite Records success story at *http://www-3.ibm.com/software/success/cssdb.nsf/CS/NAVO-5C6TVN?OpenDocument&Site=dmmain*

DB2 DataPropagator is a replication engine supporting point-in-time and near real-time replication with embedded transformation capabilities for populating data warehouses and data marts running on IBM @server® zSeries® and iSeries®. IBM IMS™ DataPropagator extends replication to IMS™ systems. These can be used with DB2 Information Integrator for mixed database replication scenarios. One IBM customer, s.Oliver, estimated that the lack of inventory visibility was costing the company approximately four percent in potential sales. It uses DataPropagator to replicate order and inventory information between staging and production servers. According to Jose Monteagudo, CIO at s.Oliver, "Although technology changes rapidly, we cannot afford to chase every fad and fashion in e-business. The software and hardware infrastructure we've built with IBM has given us the availability, scalability and reliability we can count on to help ensure a fair return on our current and future e-business investments."[5]

DB2 Content Manager enables users to integrate all forms of content—documents, Web, images and rich media—across diverse business processes and applications, including Siebel, PeopleSoft and SAP. Businesses can therefore leverage a common infrastructure, achieve a lower cost of ownership and deliver powerful new information and services to customers, partners and employees, where and when needed. As Pedro Duarte Meira, CTO and vice president of Capital IT Consulting S.A., notes, "The latest release of IBM DB2 Content Manager is a technologically advanced content management solution that offers great flexibility with new workflow solutions, unparalleled enterprise scalability, a robust and open framework and a powerful enhanced data model."

[5] s.Oliver success story at *http://www-3.ibm.com/software/success/cssdb.nsf/CS/NAVO-4ZUQ2S?OpenDocument&Site=dmmain*

DB2 Warehouse Manager rounds out IBM's information integration portfolio in the area of extract, transform and load technology. According to Sherilyn Jensen, data warehouse manager, Hillman Group, "DB2 and DB2 Warehouse Manager have proven to be a powerful combination in helping us get the most value from our data, enabling us to recoup our investment within two years."[6]

The DB2 information integration portfolio complements the WebSphere integration portfolio. DB2 Information Integrator fits easily into a service-oriented architecture, making all the power of DB2 Information Integrator available through a Web service. Therefore, any SQL operation or stored procedure operating over diverse and distributed data can return an XML result through a Web service request. Built-in tooling with IBM WebSphere Studio automatically converts an SQL operation into a Web services description language (WSDL) operation.

With IBM WebSphere Portal and DB2 Information Integrator, portal developers can reduce coding and skill requirements, speed development and incorporate more data more easily.

DB2 Information Integrator can reduce the cost and speed development of complex WebSphere business integration processes by simplifying development and maintenance of collaborations that integrate multiple data sources. And, it can improve the performance of process activities or business objects that can be modeled as sets and benefit from set-based relational processing. DB2 Information Integrator also facilitates business activity management by correlating event information with related real-time data from production systems or historical data from warehouses that may be required for an effective response.

[6] Hillman Group success story at *http://www-3.ibm.com/software/success/cssdb.nsf/CS/NAVO-4X323B?OpenDocument&Site=default*

With WebSphere MQ deployments, DB2 Information Integrator gives database programmers access to message queues using the familiar SQL paradigm. So, it simplifies integration between database and messaging systems and enables virtual or physical queue snapshots to be analyzed using standard analytical software. DB2 Information Integrator provides an easy way to compose and publish XML documents generated from diverse data sources. A single query can access diverse and distributed data, validate the document against a DTD or XML schema, and publish it to a message queue.

Finally, IBM has focused on creating a partner ecosystem to enhance analysis and reporting capabilities, enable rapid application development, enrich function delivery, extend access to additional data sources and expedite solution delivery. DB2 Information Integrator can be used with industry-leading analytical and reporting tools from Business Objects, Brio, Cognos, Crystal Decisions, Microstrategy, SAS and more. ISV partners have validated that the process of accessing diverse data is completely transparent to the user, so DB2 Information Integrator can be added to existing tool deployments without retraining. IBM has completed extensive testing with both WebSphere Studio and Microsoft Visual Studio.Net to ensure that the productivity benefits described previously apply to both Java and .Net developers. Beta testing has also confirmed applicability to the PowerBuilder environment. ISV partners also enrich feature capabilities such as customer data integration. DB2 Information Integrator can call Acxiom AbiliTec® for real-time reconciliation of customer records. IBM also partners with Ascential and Informatica to deliver the complementary ETL technology that provides data profiling, transformation, integration and consolidation capabilities. Cross Access, Information Builders Incorporated, Neon Systems and Striva Corporation extend DB2 Information Integrator reach to classic legacy sources such as Adabas, IMS, VSAM, CA-IDMS and CA/Datacom.

Summary

Businesses today need to integrate information to drive customer loyalty and satisfaction, improve operational efficiency, compete for online customers and trading partners, and identify and respond to emerging opportunities. In short, information integration provides a competitive advantage, and is fundamental to on demand computing.

A complete integration architecture should include multiple technology approaches to address the range of integration challenges businesses face. IBM understands this, providing the most comprehensive business integration solution on the market. And IBM has heard and understands the requirements for integrating diverse data. As IT Analyst Phil Howard puts it "…Oracle has preferred to concentrate upon centralization. However, the downside of centralization is that you have to rip out and replace existing databases, with all the pain that entails. Microsoft, meanwhile, has relatively limited support for federated databases in SQL Server 2000 and, even then, it tends to be limited to SQL Server support, whereas IBM has taken a more agnostic approach, supporting all sorts of relational databases within a federation. It is not hard to say, therefore, that IBM is the market leader in this space."[7] Indeed, with its new DB2 Information Integrator portfolio, IBM continues to drive state-of-the-art innovation in technologies that enable businesses to take advantage of all of their information assets.

For more information

Please contact your IBM marketing representative or an IBM Business Partner, or call 1-800 IBM CALL within the U.S. Also, visit our Web site at:
ibm.com/software/data/integration

[7] Phil Howard, "What the Hell is IBM Information Integrator?" *The Register*, February 2003.

IBM Corporation
Silicon Valley Laboratory
555 Bailey Avenue
San Jose, CA 95141
U.S.A.

Printed in the United States of America
10-03

GC18-7560-01