

Delivering information you can trust

April 2008

IBM **Information Management** software



Protect and integrate with confidence

*Continuous information protection and integration for
IBM System i*

Contents

- 2 Executive overview**
- 2 Creating a balance: integrating & protecting system i systems**
- 3 Minimizing impact – challenges and implications**
- 5 General alternatives**
- 8 The IBM® InfoSphere™ Change Data Capture difference**
- 10 The IBM InfoSphere Change Data Capture advantage**
- 11 Conclusion**

Executive overview

In today's competitive business environment, continuous system availability and data integration are key imperatives for business success. The implications of unavailable business systems and lack of integration are catastrophic. Imagine the impact on a logistics company if the system that runs its dispatch application is down. Or, the impact on a retailer if its e-business application indicates that available items are out of stock.

In pursuit of continuous and integrated business operations, more and more organizations are turning to powerful, reliable, and high performance systems to run their mission-critical operations. Today, more than 200,000 businesses in over 100 countries have come to depend on the stability and security of IBM System i™ servers. This premier server, legendary for its reliability and scalability, is used to support a business' most demanding operations. System i servers not only handle large workloads and data volumes, but are also commonly used to run the world's most critical business applications, including MAPICS, J.D. Edwards, BPCS, ADP, Geac, SilverLake, and many others.

With the widespread deployment and broad use of System i servers, it is no surprise that System i is used to power large banks in America, hospitals in Germany, manufacturing sites in China, and retail stores in England.

Creating a balance: integrating & protecting system i systems

With the importance of System i applications and the critical nature of the business operations they run, it is imperative to protect the security and scalability of the System i server.

Beyond protection, the global organizations that run System i recognize the need to analyze, share, and report on the information stored within System i applications. For example, a company's management team may want to analyze manufacturing data using business intelligence analytics and reporting tools. They may also want to share order status and production information with customers over the Internet, or monitor and report on financial and risk information for regulatory compliance. To achieve these objectives, businesses need a data integration solution that can access data on System i and integrate it into other systems and applications.

Because of the critical nature of System i applications, organizations need to deploy a solution that balances both integration and protection, while placing minimal impact on the System i environment. The ability to analyze, share, and report on business information adds value to the daily running of an organization, but it is equally important to ensure the continuous accuracy and availability of critical applications.

Minimizing impact – challenges and implications

IT professionals are under increasing pressure to integrate, protect, and recover critical business information in less time and at a lower cost. This requirement is challenging to meet unless low-impact, scalable solutions that work with existing IT infrastructure can be found. Unless careful consideration is given to how selected solutions operate; there can be adverse effects on system operations, including:

- *Delayed processing times*
- *High network bandwidth usage*
- *Slow response times for critical operational tasks*
- *Security risks*
- *Additional demand for batch windows, required for processin*

The implications of placing these demands on system resources are detrimental to the operational performance of critical System i applications. With system performance hindered, users are prevented from accessing critical business applications from System i systems, therefore interrupting a business' ability to run normal operations.

Therefore, when searching for a protection and integration solution for a System i environment, it is important to consider a solution that imposes minimal overhead on system resources, ensuring the efficient and timely access of critical System i applications.

The following are some of the factors that should be considered when selecting a solution for protection or integration:

- *Does the solution require changes to be made to applications? If yes, how will the changes be tested before they are put into production?*
- *Does the solution require triggers or other changes to existing data models?*
- *Does the solution need to be modified when new applications are added or data models changed?*
- *Do new tables or records need to be created in the database? If new tables are created, how large will they get?*
- *How much disk space does the solution use?*
- *Does the solution have the ability to monitor its impact on the systems?*
- *How many database and network resources are required to transfer data between the integrated systems?*
- *Will the integration solution cause disruption to users or require critical applications to go offline in order to capture the data?*
- *How large are batch windows for data extraction?*

All of these factors need to be taken into account when assessing the flexibility and scalability of a solution. Integration and protection solutions that support the low impact/high performance paradigm are recommended because they require minimal change to existing applications and data models, enabling rapid implementation and quicker deployment. Solutions that follow this paradigm also use little network bandwidth, CPU, and database resources, resulting in rapid, uninterrupted response times for critical operational systems.

By efficiently integrating data between operational systems, users can make critical business decisions in a timely and efficient manner.

General alternatives

Whether for protection or integration, there are a number ways of extracting data from operational database systems. However, each approach varies significantly in terms of performance, scalability, and effectiveness. Because of the critical nature of applications stored in System i environments, it is important to select an integration method that best supports and secures an existing IT infrastructure and reduces the impact on the network.

This section addresses three common methods and explores the impact each can have on the System i environment.

- 1. Extract, Transform, and Load: Traditional Extract, Transform, and Load (ETL) tools are widely deployed by organizations performing complex transformations on large data volumes, usually in support of data warehousing initiatives. ETL tools require additional functionalities to handle continuous, real-time information requirements.*

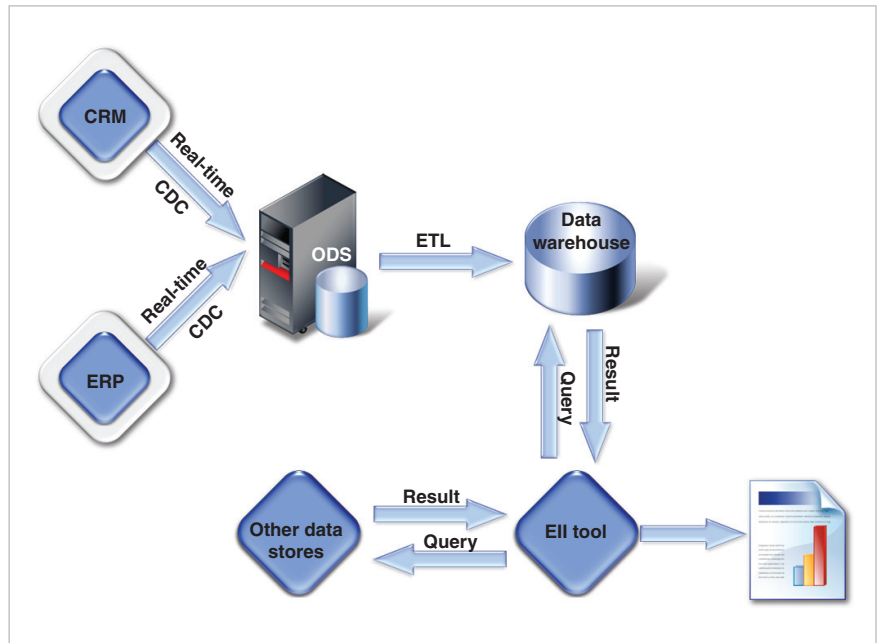
2. *Enterprise Information Integration: Another common method used by organizations to integrate data from System i is by executing queries directly against the database. This method is often called Enterprise Information Integration (EII). Every time this method is executed, a query is sent to all defined source systems. With additional users, more queries are run against operational System i application. This solution requires direct access to source data.*
3. *Change Data Capture: As the sources of data; and the volume contained in those sources continue to increase; and the pace of business accelerates, real-time change data capture (CDC) technology is becoming more widely deployed within System i environments. CDC solutions enable data protection and integration in a scalable and minimal impact manner.*

CDC technology offers a low-impact alternative for data integration because it eliminates redundant data transfer and reduces network bandwidth. This is achieved by capturing changes made to data in the source application and sending only the changed data to target applications. By replicating data directly from the source to the target, CDC solutions eliminate the onset of risky changes—such as the addition of servers or programming changes to applications or data models—to an organization's existing IT infrastructure.

Real-time CDC technology can be used in conjunction with ETL and query-based solutions to make integration more frequent, selective, and non-intrusive. Used in combination with existing IT investments such as ETL or EII tools, CDC technology allows organizations to gain real-time awareness of production systems.

The following diagram depicts the complementary use of CDC technology with ETL and query-based or EII solutions while also meeting protection requirements.

Figure 1: CDC technology with ETL and query-based or EII solutions



Each of the above approaches delivers value in one form or another. To protect the critical nature of System i systems, it is important to select a solution, or a combination of solutions, that assures the high scalability and security of data, while ensuring the highest level of performance for IT environment.

The IBM InfoSphere Change Data Capture difference

IBM InfoSphere Change Data Capture's real-time CDC technology includes a flexible architecture that allows organizations to select the best integration and protection option for their needs. It is important to select a solution that provides different alternatives because the solution must have the ability to adapt to changing IT needs – whether those are application or system upgrades, spikes in data volume, or increased user access.

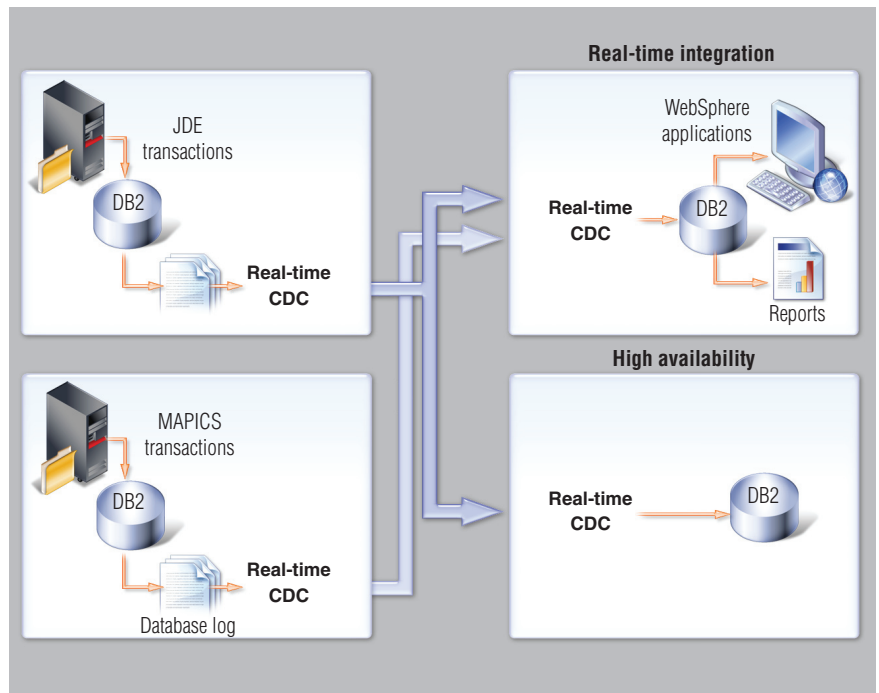
IBM's Integration Suite is a complete data integration and protection solution that supports a wide range of systems and applications, including DB2® on System i, System p, and System z, to enable efficient load balancing, e-Commerce activity, reporting, and other critical functions. The integration suite enables real-time CDC through its continuous database log-scraping architecture that moves data directly between source and target systems. As soon as changes are detected in the source system they are immediately transformed and flowed to one or more target systems. Businesses experience rapid system response times by offloading resource-intensive events from the primary System i.

IBM InfoSphere Change Data Capture also supports Web environments through event triggering, load balancing, or through the continuous feeding of message queues. This support allows WebSphere® applications to gain broad access to data without compromising the integrity or performance of production applications on the primary System i system. Available from the same vendor; IBM InfoSphere Change Data Capture complements WebSphere Information Integrator software by providing the ability to create a real-time operational data store that can be accessed by Information Integrator or other ETL tools. Users can get real-time integrated data without having to directly access – and therefore impact – critical System i applications. The data collected within the real-time ODS can be used as a source for Information Integrator to combine with information from other data stores. With effective workload balancing and data sharing, businesses can increase operational efficiencies and reduce network bandwidth and CPU usage.

In addition to data integration, CDC technology also supports high availability and business continuity within System i environments by ensuring that all systems remain up-and-running.

The following diagram depicts the complementary use of InfoSphere Change Data Capture and high availability solutions.

Figure 2: InfoSphere Change Data Capture and high availability solutions



The IBM InfoSphere Change Data Capture advantage

Building on years of System i integration and protection experience, IBM InfoSphere Change Data Capture provides customers in all industries with effective real-time solutions. Most, if not all System i customers depend on their systems to run critical MAPICS, Enterprise One, BPCS, SAP, Jack Henry, SilverLake, and other business applications. These same customers rely on IBM's technology to gain continuous, secure access to data within these System i enterprise applications so they can conduct efficient e-business, reporting, business analytics, and other business functions, without impacting the operational performance of their applications or slowing down their business.

The following are just some of the business benefits that can be derived from IBM's real-time CDC technology:

- *Low Impact on Operational System Performance – Data is accessed directly from the database logs rather than by performing queries against the database, ensuring the peak performance of operational applications and systems. Performance is also maintained by spreading the loading process of the DB2 database over time, thereby avoiding batch windows.*
- *Non-Intrusive Architecture – Minimal overhead on IT architectures by detecting all database changes at the transaction level through the parsing of the database log, eliminating the need for intrusive timestamps, database triggers or changes to the application. Peer-to-peer architecture does not require additional hardware purchases for proprietary staging stores or metadata repositories and transfers data directly, without the need for additional third-party gateway technology.*
- *Low Data Latency – Because data is captured and integrated as changes occur on source applications, data remains up-to date, allowing businesses to make accurate decisions that reflect the current state of their business.*

- *Increased Scalability – Peer-to-Peer architecture eliminates central server bottlenecks, allowing data to move efficiently throughout the organization. By continuously moving changed data only, it can be scaled to meet wide-ranging data volumes. Because data is integrated in real time, organizations are not required to fit their data integration processing needs into shorter batch windows; the integration process can occur throughout the day without interrupting daily operations.*
- *Real-Time Data Warehousing – By capturing before-and-after images of source database updates, these can be used to incrementally update dimension and fact tables in data warehouses on the target system. This allows users to perform analysis on changing data vs. static data (e.g. analysis on orders in progress vs. completed orders).*
- *Rapid Implementation – The solution is easy to set-up, configure, and manage. This ease of use shortens time-to-production, delivering more value for less time and money.*

Conclusion

This whitepaper has examined common solutions for System i integration and protection and has provided insights into the features that an effective solution should employ to reduce impact on a System i environment. With a low-impact, high-performance data integration solution from the same vendor as System i's; users can effectively analyze, share, and report on the information stored within mission critical System i applications and avoid the implications that can result from degraded performance of operational systems. With continuous access to accurate information stored in critical System i applications, organizations can make better business decisions that allow them to take timely action and move forward faster.

For more information

For more information about IBM Information Server, contact your IBM marketing representative or visit ibm.com/software/data/integration



© Copyright IBM Corporation 2008

IBM Software Group
Route 100
Somers, NY 10589

Printed in the United States
April 2008
All Rights Reserved.

IBM and the IBM logo, InfoSphere, DataMirror, iCluster, System i, DB2, WebSphere and i5/OS are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries or both.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc., 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.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

TAKE BACK CONTROL WITH **Information Management**