

Easy, Automated, Real-Time Data Sharing: What Can It Do for Your Business?

**Smarter Data =
Smarter Decisions**

WHITEPAPER



Data Sharing: Top Management Issues

The demands on a business today – increased global competition, lean budgets, big data, increased M&A activity and tighter regulations – are creating an ever-increasing need for access to data. In this environment the accelerating “time-to action” will become the new business precept. At its core is the ability for every organization to ensure that the right information reaches the right people at the right time. The challenge: Taking the modern landslide of data and extracting from it real business intelligence and value.

What if you could effortlessly replicate data from virtually any popular database into another - and back again - without custom programming or large-scale integration costs?

What if you could share data from one application to another in real-time and be assured that the data would be in the exact format each application and user needed? Would it help if you could add a new customer, update an inventory balance or delete a stock item, and have all of these applications share the same information in real-time?

What would your business gain if you could break down isolated silos of user data and give everyone across the enterprise the same effortless access to accurate, current and complete data?

What if you could accomplish this without discarding or changing your investment in existing systems and databases? How much more productive could your database staff be if a single solution could be used to manage data sharing across all databases across all platforms?

Would it help you to handle data security, auditing, and compliance with government mandated regulations using a standardized, enterprise-wide tool?

What if you could be confident that the right data is reaching the right people so you could spend your time on other pressing IT/business alignment issues?

Do These Situations Sound Familiar?

- **Integrate Multiple Organizations.** A company recently merged with a rival. Their centralized applications have used the same standards for years and are consolidated onto an IBM® Power Systems platform running AIX® and Oracle/ SAP®, along with several Windows® servers. The newly acquired company relied on Power Systems server DB2® databases to run critical applications. How do they integrate these organizations without undertaking a costly, year-long integration effort?
- **Integrate Heterogeneous Platforms.** An end user community needs real-time query access to selected Power Systems server production data kept in a DB2 database. But they would like this data rendered to a MS SQL Server database that hosts GUIs and applications that they are familiar with.
- **Push or Pull Data in Real Time.** A distributor of health care products is deploying an eBusiness web portal using Windows/SQL Server to allow customers to track and update orders. The application needs to pull and push data real-time to the backend JDE ERP tool running DB2 on the Power Systems server.

“The one word that best describes the state of analytic data in large organizations is “fragmented.” Despite their best intentions, CIOs are struggling to deliver consistent data that provides a single view across the enterprise.”

—TDWI

- **Integrate Disparate Databases.** A general ledger system using MS SQL Server databases needs current data from the HR payroll system running on Oracle, as well as from the MS SQL Server-based accounts payable, billing, and accounts receivable applications, and also the manufacturing supply chain application using Sybase®.
- **Eliminate Custom Programming.** The marketing and sales departments use an application that was never designed to share information with the new CRM and ERP applications. The custom programming required to provide real-time data sharing between and among these applications will require a costly budget increase to bring in outside consultants.
- **Accommodate New and Advanced Systems.** Healthcare personnel need to have access to patient information records that reside in a DB2 database on IBM AIX. They are using an application that deploys multiple laptops across the hospital floor using a back-end Windows SQL Server database to manage and record daily patient care details.
- **Establish Centralized Data Warehouses.** Your company invested in a high-end Teradata® data warehouse. You need to feed data into this centralized data store from several other databases and systems in your organization, which include AIX and Windows running DB2 and MS SQL Server. Your executives expect extensive dashboard access with real-time updates.

The Technical Challenge

In today's IT environment, critical business information often resides across a variety of operating systems and databases within the enterprise. The reasons leading to this configuration are many, but the result is often silos of information that cannot be shared with other silos. Frequently, portions of the data contained in each of these silos are redundant, but not necessarily in-sync across platforms. This puts the business at risk of making decisions based upon divergent, inaccurate data sources. Isolated information silos simply do not work well in the real-time, on-demand business world.

Populating data across different databases has traditionally been done with ad hoc manual processing. This requires DBAs and application owners to develop and maintain customized procedures, commonly referred to as ETL processing:

- **Extract:** Pull the data from the source database into a flat file
- **Transform:** Perform data transformations required for proper input to the target database
- **Load:** Load the data from the flat file into the target database. (Note: this step often also includes copying the flat file over to a separate, potentially remote, target system before loading.)

“IT is continuously challenged to find easier ways to share data between application silos without spending a lot of money, without re-hosting applications or performing significant application surgery, and with real-time data sharing.”

— DM Review

“One of the curious epiphanies that strike users as they study the economics of data integration is that the cost of hand-coding is misleadingly low in the first phase of a project, but it goes through the roof in later maintenance phases.”

— Forrester Research

The Total Cost of Ownership of ETL

In a large enterprise, numerous ETL processing events can occur weekly, all of which must be scheduled and monitored. Since ETL processing is often done daily during off-hours, data concurrency among the target databases may still lag by as much as 24 hours. Also, there is usually no centralized means to identify all of the ETL scripts and processing that occur across an enterprise, so keeping track of scripts and updating them as needed can be a tedious task, and is prone to oversight.

Custom, home-grown scripts are costly to develop and maintain. As databases grow, upgrade, or change, more work is required to link them using traditional integration techniques, especially one-to-one programs or scripts. This solution is difficult to scale in a dynamic environment.

Replication products from database vendors represent an alternative to ETL script processing. Many products exist for replication, however these typically support only like-to-like source and targets, and most do not offer database data extraction or transformation features. Most RDBMS vendors offer replication products, but to add support for heterogeneous databases, the solution may need an add-on gateway product. Often these vendor-provided solutions are difficult to manage and may not even be able to support replication between different releases of the same product, so their scope is very specialized and very limited.

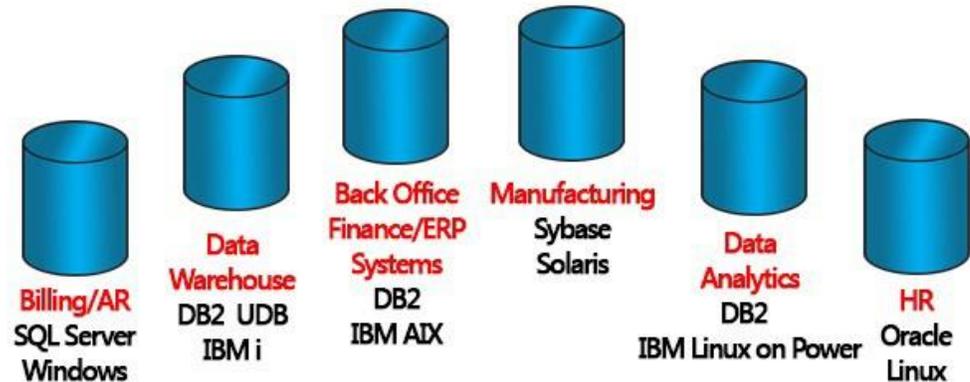
Another approach often considered is use of an application integration product or message oriented middleware (such as MQ Series™) between databases. This provides a general application framework to send and receive data, but requires significant application rework and additional staff expertise to support the added messaging layer, which can be a complex and significant investment.

Make versus Buy: The Quest for Return on Investment

IT has come under increasing scrutiny because of the costs of data integration projects. Any ROI calculation must account for the following costs, using a fully burdened labor rate.

- **Development Costs:** Each data-sharing source and target must be treated as a separate project. As requirements change each source and target pair will become a separate development project.
- **Cost for Test Systems:** These include RDBMS licensing, system resources, storage, and additional RDBMS add-on products such as gateways that may be needed to implement an internally developed solution.
- **Maintenance Costs:** These costs include maintaining an internally developed data sharing solution that requires ongoing coding changes to adjust to new or altered data schemas, along with monitoring the running and accuracy of the supporting processes and testing with new DBMS/OS releases. Turnover in your IT department, which results in fewer people who are fully familiar with your data and application structures, can dramatically increase the costs of all of these activities.
- **Opportunity costs.** There are costs in missed opportunities when IT staff has been dedicated to developing and maintaining an internal replication solution versus working on other internal projects.

Barriers to Information Sharing Isolated Corporate Data Silos



A mixed environment like this adds complexity and creates data sharing and data synchronization challenges for IT professionals.

The Alternative: Real-Time, Automated Data Sharing

Acquiring a standardized data sharing solution fits your needs if you want to quickly solve pressing data integration issues between critical applications that use databases such as Microsoft SQL Server, IBM DB2, Oracle and Oracle RAC, Teradata, Informix and Sybase. A standardized solution can provide access across the enterprise to accurate, real-time data, regardless of where it is created or resides, while saving you the cost of expensive and complex data integration projects.

A standardized software-based data sharing/replication solution can work well in your environment if you want to spend a minimum amount of time setting up and managing the data sharing process. This approach saves time and money over data sharing tools that are developed and maintained internally.

Meet the Challenges with CloudReplica dbShare

CloudReplica dbShare is a robust yet flexible solution for data sharing. It gives you a simple, on-demand way to share data across your enterprise, leveraging every piece of data in your organization and making it available in real-time to all of your applications. It ensures high data quality and zero data loss. And it is standardized, meaning that it uses the same management interface to define the extraction, transformation, and replication processes across all database types, across any operating system platform.

The CloudReplica dbShare GUI management console enables administrators to define replication models in a familiar user interface. This method is much more intuitive and expeditious than programming each replication model by hand and allows you to implement a standardized, table-based approach for data sharing/replication. Furthermore, the tool approach allows a model be quickly and easily changed or expanded to include other databases as your databases grow or change.

“The need for ever fresher information is accelerating... this in turn puts increasing pressure on data integration to operate faster and more frequently so that information freshness is measured in hours or minutes instead of days.”

— ForresterResearch

This flexible modeling approach also enables to you to deliver information bi-directionally between source and target databases, without disrupting normal operations. For example, you can quickly and transparently enable information distribution from a central office to remote locations, with just a few clicks. Updates from the remote offices can then be immediately applied back upstream to the centralized database for designated transactions.

CloudReplica dbShare captures data changes as they happen, so it can replicate data between databases as easily in real-time as on a scheduled basis. Advanced replication options include broadcast and cascade configurations.

CloudReplica dbShare works with most RDBMSs in the marketplace, including Microsoft SQL Server, IBM DB2, Oracle and Oracle RAC, Teradata, Informix and Sybase. CloudReplica dbShare provides administrators with a single control point for managing data sharing across all databases, and across popular operating systems. For example, you can set up CloudReplica dbShare to transform Oracle data into the exact format needed for DB2 or MS SQL Server databases and applications. Simultaneously, it can transform DB2 and MS SQL Server data into Oracle data types. And when moving data between mainframe and open systems databases, it even performs the required EBCDIC/ASCII conversion automatically, while the data is in flight.



Data Sharing: Oracle and Geo-Spatial Data

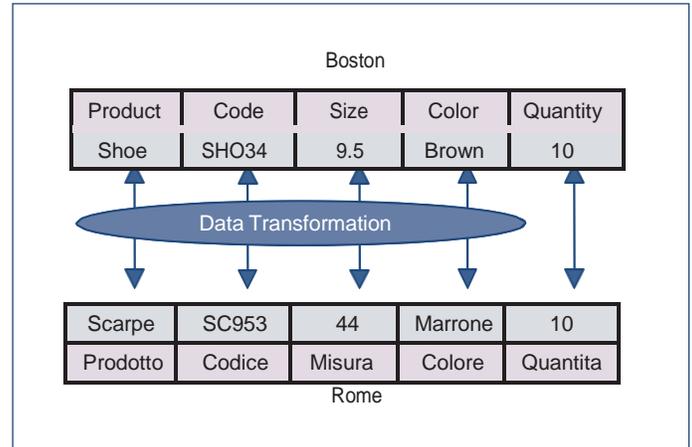
Challenge: To make continuous, real-time data available to its stakeholders via the internet, this US-based land management company, serving large and small non-industrial private forest- land owners, investors, conservation organizations and forest industry businesses, needed an information availability tool that could cost-effectively and efficiently deliver data where, when and how users needed it. Its use of large-format, geo-spatial data posed a significant integration challenge. The company needed to make this data easily available to customer applications as well as to multiple databases, like Oracle, within its own business groups.

“We were able to see a high level of productivity and greater IT efficiency with CloudReplica dbShare because it was able to handle our large geo-spatial files easily without programming.”

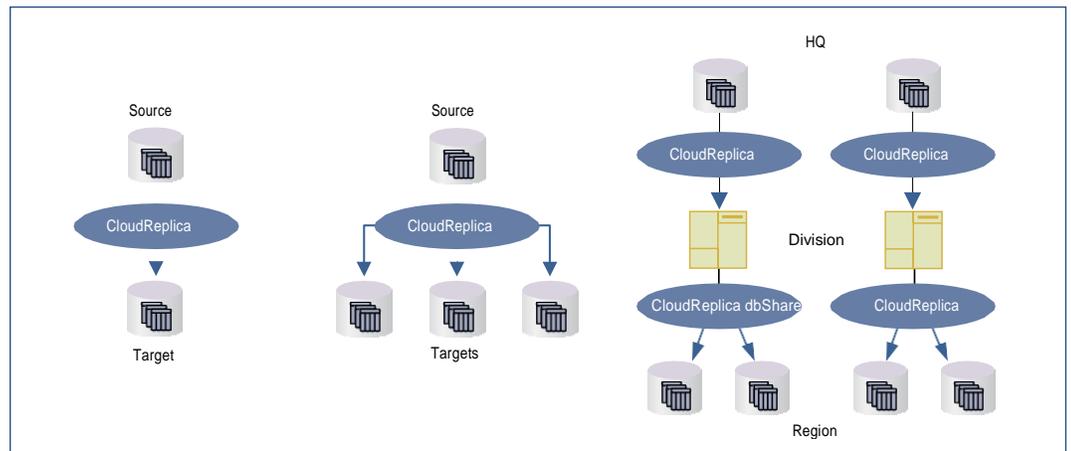
Solution: The company realized that only CloudReplica dbShare was robust enough to take on its critical geo-spatial data and make it available to customer and user applications. The company was attracted to the solution’s ability to drive all of the data sharing processes at the database level, while the management console made installation easy. CloudReplica dbShare’s table-based structure easily adapted to accommodate the unique needs of the client’s internally developed application and its business processes without custom programming, saving time and money. The built-in automation saved administrative effort and could be quickly customized and managed with only a few clicks.

CloudReplica dbShare Features

- Robust operations console minimizes DBA tasks and monitoring chores, which lowers the total cost of ownership.
- Set it and forget it automatically notifies administrator of any errors via email.
- Model-based set up and configuration works at the database level with total application independence. CloudReplica dbShare includes extensive defaults so installation can be as simple as just a few clicks.
- Business rules-driven—over 60 pre-built, click-and-go data transformation functions are included to eliminate setup chores. Customization tools provide further options and control. For example, if a formerly independent business unit needs to share information with your finance department, you can easily change the business policies using the CloudReplica dbShare console, setting up the appropriate replication processes for a one-time event, or for ongoing data sharing.



The business-rules based tables ensure data is transformed as needed and reaches targets in the required format and language without custom programming.



Flexible replication methods can solve data challenges regardless of source or target database. For example, the “Paths” mode allows you to specify how data will reach targets using intermediate hosts, which do not require a resident database server. This supports cascade replication as well as replication through firewalls and gateways.

- Flexibility—leverages Java-based, open standards technology to quickly and inexpensively support new platforms and operating systems while maximizing ROI. CloudReplica dbShare uses JDBC database connectivity and XML, eliminating the cost of building and maintaining code unique to each database and writing one-to-one scripts, and ensuring adaptability to future changes in database technology.

- **Audit Journal mapping**—captures the details of any update activity occurring on a table and records them on a target table (journal) or flat file for auditing purposes. This standard process helps you meet Sarbanes-Oxley Section 404 compliance regulations automatically by establishing a data trail when financial information changes, including before-and-after values of every column, the type of transaction, and the sending DBMS type, table name, user name, and transaction information.
- **Data integrity**—ensures that only committed transactions on the source replicate to target(s) and are applied in the order they were processed on the source. CloudReplica dbShare resolves any conflicts when two updates hit simultaneously, ensuring that all sites are kept in sync for complete data consistency. CloudReplica dbShare automatically re-applies any uncompleted transactions, should any interruption occur. The result is assured zero data loss, which means consistently reliable information for users and executives.
- **Modeling tools**—enable you to test changes to your replication model before you implement them. This acts as a safety net to prevent any potential problems from affecting production databases. This feature can automatically detect and fix any errors as well. Not only does CloudReplica dbShare validate models before they are committed to production, it also provides automatic versioning to assist in the control of what is in production.

CloudReplica dbShare Expanded Roles: A Powerful, Multi-purpose Data Tool

In addition to its primary purpose of providing real-time data sharing, CloudReplica dbShare has also emerged as a tool that can play more than just one role in your organization. Below are some examples of different ways CloudReplica dbShare can be leveraged in your environment:

Smooth integration challenges:

CloudReplica dbShare gives you the ability to efficiently integrate systems after a merger or acquisition by eliminating the need to immediately migrate and convert to a common database or operating system. It provides a fast and easy solution to integrate applications that are acquired, without the need to dig into the detailed code to make extensive changes right away.

Data Sharing: IBM DB2 and Oracle Data

Challenge: A South American-based cable television provider, serving more than 1.5 million subscribers, faced the daunting challenge of a heterogeneous IT environment. Its operational systems run on IBM® Power Systems™ servers and DB2 Databases, but its marketing, sales and web applications run on Sun® and Oracle®. Sharing data between these differing technologies was difficult and made marketing analysis a heavy drain on computing resources. Consequently, the company was reluctant to provide marketers with direct access to operational data.

“The installation was quick and painless . . . we estimated that it would have taken three months without CloudReplica dbShare.”

Solution: CloudReplica dbShare now simply and transparently integrates information across the different DBMS, operating systems and hardware platforms. The company easily replicates data between any source and an LPAR-based target and between the Power Systems server and Sun platforms. Marketing users now have on-demand access to data for analysis and other purposes via the replica without impacting production performance. IT staff using the Sun/Oracle platform gain easy access to data that originates in DB2 on

“People trying to compete in the global marketplace,” ... “Can you squeeze your lead times? Can you tailor product for specific customer needs? You need real-time answers in order to serve your customers.”

— Power Systems Network

- **Address corporate governance information protection and access requirements:** CloudReplica dbShare auditing capabilities help executives understand what happened to data. The journal itself can be saved for as long as required, so it can provide an archived audit trail. Also, queries can be run against the journal archive for ad hoc questions arise, without affecting the production database. The fast real-time translation and replication features mean you can also move disparate data to one repository in one format to satisfy HIPAA regulations as well.
- **Support data warehouse strategies quickly, effectively:** The table-based structure can be adapted to fulfill data warehousing needs, including moving information from multiple locations to a standard format within a central warehouse, keeping the warehouse up-to-date with real-time datafeeds.
- **Solve the “make versus buy” dilemma with a cost-effective solution.** The table-based structure and non-application specific nature means there are no costly, time-intensive integration projects. The modeling functions enable you to test new data sharing strategies quickly, or make changes to existing strategies, without impacting users or production.
- **Offload production.** CloudReplica dbShare can replicate real-time data from production to a second database located on a “reporting” server. Business users can then run real-time queries and analytics against the “reporting” target, keeping ad hoc queries off of the production copy of the database, avoiding performance impact.
- **Database migration.** CloudReplica dbShare enables real-time replication from a current or legacy production database to a new target database running on a different host or database platform. Ongoing testing can occur on the target before cutover. Resulting cutover windows are short and minimally intrusive because your data has already been moved and tested on the new target.
- **Save disk space.** The need to extract and FTP or copy files, as part of “legacy” ETL processes, frequently requires large “holding areas” on disk to hold files prior to loading. This is no longer needed with CloudReplica dbShare, instantly giving you back valuable storage capacity.
- **Make better use of DBAs.** Multiple, manually developed “legacy” ETL processes require significant time to implement, track, monitor, repair, and verify. This effort is eliminated with a single, standardized tool for implementing database data sharing.

Smarter Data = Smarter Decisions

As businesses strive to be more effective and efficient, there is a greater need to have access to accurate, real-time information for making informed business decisions. Real-time access to information is now seen as a necessary competitive advantage. Without it, businesses large and small face greater risks from decisions made with data that is less than 100% current and accurate. Replacing legacy processes with dynamic data sharing makes your data smarter, resulting in a smarter business model.

CloudReplica dbShare provides a central, standardized, easy-to-manage, click-and-go capability for sharing data between disparate databases and application silos, without programming changes or integration projects. CloudReplica dbShare fully leverages the flexibility of open standards technology into your database sharing strategy. This results in fast adaptability to IT changes and future growth plans and suits companies that need a cost-effective solution with a fast ROI.

About CloudReplica

With headquarters in Houston Texas, CloudReplica offers customers 'business continuity' by eliminating day-to-day backup management and automatically sending data changes as they occur to a secure off-site cloud service. Your company deserves real-time business protection.

For over 7 years' companies have relied on CloudReplica to develop, deliver, and maintain true Enterprise Class Business Continuity solutions to protect their most critical business processes.

The CloudReplica executive team has a long history of developing and deploying Disaster Recovery and Business Continuity solutions for some of the world's largest Fortune 500 companies such as Exxon, Boeing, Chevron, Hewlett-Packard, and BP Amoco.

You can be confident that the professional team standing behind CloudReplica will deliver you World Class support and be there at the critical time when you need them.

Bob Reiter
Business Development Manager
US and Canada
Bob.Reiter@CloudReplica.com
+1 832.628.9041
+1 800 374 3220



www.cloudreplica.com

