

IBM InfoSphere Data Replication

Highlights

- Provides dynamic, near-real-time, delivery of transactional relational data
 - Delivers data from IBM Db2, Oracle, SQL Server, PostgreSQL and more
 - Integrates real-time inventory and sales information
 - Enables integration of product and customer data
 - Supports 24x7 worldwide operations with reliable, exceptional performance
 - Provides the data synchronization for continuous data availability
-

Provide near real-time incremental delivery of transactional data

To compete in today's fast-paced business climate, enterprises need accurate and frequent sales and customer reports to make real-time operational decisions about pricing, merchandising and inventory management. They also require greater agility to respond to business events as they happen, and more visibility into business activities so information and systems are optimized for peak efficiency and performance. By making use of data capture and business intelligence to integrate and apply data across the enterprise, organizations can capitalize on emerging opportunities and build a competitive advantage.

The IBM InfoSphere Data Replication portfolio is designed to address these issues through a highly flexible one-stop shop for high-volume, robust, secure information replication across heterogeneous data stores.

The portfolio leverages real-time data replication to support high availability, database migration, application consolidation, dynamic warehousing, master data management (MDM), service-oriented architecture (SOA), business analytics, and extract-transform-load (ETL) or data quality processes. It also delivers outstanding capabilities for loading real-time information into a data warehouse or operational data store, which can help organizations enhance business agility and visibility into key processes.

In addition to direct delivery of changing data to Apache Hadoop, IBM InfoSphere Data Replication can deliver data to Apache Kafka, providing an ideal big data landing zone and point of enterprise integration for changing transactional source data. Using a Kafka landing zone simplifies the delivery of operational data to multiple consumers such as Hadoop Distributed File System (HDFS), cloud data stores and event-driven consumers.

Leverage a comprehensive solution for data replication

The data replication portfolio includes multiple technologies to support the broadest set of source and target, or both pairings as well as a comprehensive set of data replication topologies. These include:

- Change data capture (CDC)
- Q Replication
- SQL Replication

CDC in Figure 1 addresses a broad spectrum of data synchronization challenges by delivering changed data from a relational database source directly to:

- A target relational database management system (RDBMS) in support of a data warehouse, operational data store or cross-application integration
- A target Hadoop or Kafka cluster in support of enterprise information hubs and other big data initiatives
- A cloud-based version of one of the above
- An IBM PureData appliance for high-performance analytics
- IBM InfoSphere DataStage and IBM InfoSphere QualityStage for transformation, rich data integration and data quality purposes
- IBM Master Data Management Server for enterprise data integration initiatives

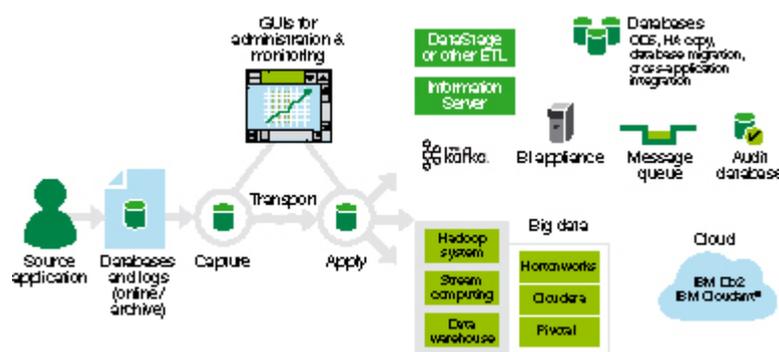


Figure 1. Log-based change data capture (CDC)



Figure 2. Q Replication architecture

Q Replication in Figure 2 offers queue-based replication by leveraging IBM WebSphere MQ for data delivery and fast recovery for high availability after a system outage. Q Replication is ideally suited to the synchronization of IBM Db2 data across the data center or around the globe.

Key specific offerings in the IBM data replication portfolio include:

- IBM InfoSphere Data Replication and all of its permutations for IBM Db2 for z/OS, IMS for z/OS, VSAM for z/OS and non-RDBMS targets
- IBM InfoSphere Data Replication and its permutation targets for IBM z/OS sources
- IBM InfoSphere Classic Change Data Capture for z/OS - IMS and VSAM
- The new IBM InfoSphere Data Replication Db2 for z/OS remote source delivers a new deployment model for the dynamic delivery and on-going synchronization of mainframe IBM Db2 for z/OS data with Linux, Unix, Windows (LUW) and cloud data platforms that minimizes z/OS MIPS used by shifting source capture processing to an LUW platform
- IBM InfoSphere Data Replication for IBM Db2 for providing high-availability for IBM Db2 Warehouse environments that are deployed on IBM Integrated Analytics Systems

Customers have the flexibility to acquire and deploy components to fit their needs.

Make decisions based on the latest, most accurate information

The IBM InfoSphere Data Replication portfolio provides real-time feeds of changed data for data warehouse or MDM projects, enabling operational and tactical business decision-making based on the latest information. It dynamically routes data for consumption by one or more applications to help ensure accurate, reliable data is used across the enterprise.

Reduce costs by facilitating application consolidation

As businesses cope with economic uncertainty by cutting costs, IBM InfoSphere Data Replication helps accelerate application consolidation projects by eliminating multiple application instances on different systems, databases and operating systems. This approach

can help reduce maintenance costs dramatically.

Help reduce risk through exceptional performance

Consolidating and moving the data contained within applications can be a sensitive task. Consolidation can be challenging because most system environments comprise a variety of operating systems and databases. Also, if consolidation or migration projects require significant downtime, it could defeat the potential cost savings generated by the initiative.

IBM InfoSphere Data Replication addresses this issue by allowing data to be migrated from one database or platform to another while end users are still using applications—without downtime or productivity loss for the business. It also supports a broad range of operating systems and databases to help maximize the value of investments for use in future projects.

Instead of using triggers or performing queries against the database, IBM InfoSphere Data Replication software reads the native database log to capture changes. For businesses faced with shrinking batch windows or overutilized applications, this log-based CDC approach helps ensure the performance of even the most demanding mission-critical applications running on the source system is not adversely affected.

In addition, data replication supports near-zero latency for pervasive integration projects.

Data replication in the real world

Data replication is a critical capability for a variety of industries. For example:

- **Financial services companies** must leverage data from legacy financial systems using open platforms that support efficient internet-based communications, statement processing and other applications with minimal impact on key systems.
- **Retailers** must have the ability to sell in stores, over the web and by phone—and they need consistent, up-to-date information that is accessible across all three channels. Data replication can help ensure, for example, that wedding registry purchases are not duplicated and that customers can check product availability and order status over the phone or the web.
- **Shipping companies** must allow customers to access real-time shipment information online, but they must also significantly reduce the impact on their own mission-critical logistics systems to support efficient, continuous operations.
- **Government agencies** need to provide public access to up-to-date information on criminal proceedings on the internet, but they cannot risk unauthorized access to sensitive information on offenders.

- **Health insurers** need to offload queries, reports and backups to avoid negatively affecting performance and user response times on essential claims processing systems.

Enable consistent, trusted data delivery

The downtime costs for the inaccessibility of mission-critical data and applications range from contractual fines and lost productivity to a loss of credibility—all of which may result in lost customers. With IBM InfoSphere Data Replication software, data can be synchronized between two systems to provide continuous availability. If the primary system is impacted by a planned or unplanned outage, a secondary system is available to serve customers and keep the business up and running.

IBM InfoSphere Data Replication enables continuous delivery of data to support critical business operations. Transaction consistency is maintained throughout the process to preserve units of work and referential integrity. The software supports full transaction granularity with before-and-after images of all transactional changes. The platform is data event-aware, so it can be used to trigger specific business processes. In addition, fault tolerance capabilities allow organizations to recover to the last committed transaction.

Why IBM?

IBM supports enterprise DataOps with data integration tools to transform structured and unstructured data and delivers it to any system.

IBM InfoSphere Data Replication solutions provide trusted data integration and synchronization to efficiently manage data growth. It powers the use of real-time information for DataOps by enriching big data systems and mobile applications, even capturing data that is constantly changing.

For more information

For more information about IBM InfoSphere Data Replication solutions, contact your IBM client representative or IBM Business Partner. Learn more at ibm.com/data-replication.

Learn more about DataOps at ibm.com/dataops. Follow us on Twitter at @IBMDData, on our blog at ibmbigdatahub.com and join the conversation #DataOps.

Additionally, IBM Global Financing provides numerous payment options to help you acquire the technology you need to grow your business. We provide full lifecycle management of IT products and services, from acquisition to disposition. For more information, visit: ibm.com/financing

© Copyright IBM Corporation 2020.

IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at <https://www.ibm.com/legal/us/en/copytrade.shtml>, and select third party trademarks that might be referenced in this document is available at https://www.ibm.com/legal/us/en/copytrade.shtml#section_4.

This document contains information pertaining to the following IBM products which are trademarks and/or registered trademarks of IBM Corporation:
IBM®, IBM PureData®, DataStage®, Db2®, DB2®, InfoSphere®, QualityStage®, WebSphere®, z/OS®



Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

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

UNIX is a registered trademark of The Open Group in the United States and other countries.

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