

Using IBM InfoSphere Change Data Capture to complement your existing IT environment

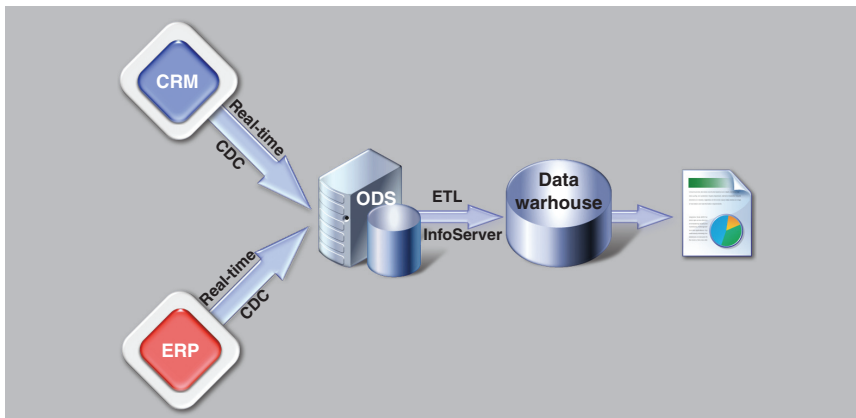


Figure 1. Architectural overview of traditional ETL tools

Highlights

- **Complement their current data integration strategy**
- **Real-time Change Data Capture technology**
- **Accesses the data directly from database logs**
- **Without requiring batch windows or interruptions to mission-critical systems**

Traditional ETL (extract, transform and load) tools were originally built to help companies load data into data warehouses and data marts for analytics and long-term forecasting. Because time wasn't of the essence, data was usually batch-loaded on a daily, weekly or monthly basis using tools that bulk-load large amounts of data and perform the complex aggregations and summarizations necessary for reporting from the data warehouse.

But today, business changes are dynamic and time-sensitive. Executives need access to real-time reports so they can make better business decisions; the sales team needs up-to-date customer data to improve cross-selling and up-selling opportunities and manufacturers need access to accurate inventory to enable just-in-time inventory. The amount of data being integrated across applications and being loaded into data warehouses is growing substantially and, conversely, the amount of time that these source applications have in order to perform batch processes is shrinking. As competition heats up, it is no longer acceptable to work with day-old or week-old data.

Organizations need a way to access data from source applications throughout the day, without requiring lengthy batch windows and without impacting the performance of mission critical operational systems.



Today, many IBM customers transform their data and load their data warehouses using ETL tools but turn to IBM's change data capture technology to complement their current data integration strategy with the real-time advantage. The value addition is the Change Data Capture technology that accesses data directly from database logs and replicates data from operational systems without requiring batch windows or interruptions to mission-critical systems. The data is then loaded into operational data stores (ODS), with the ETL tool then feeding the data warehouses or other applications.

Change Data Capture complements traditional ETL tools by:

- *Providing real-time data for operational reporting – Data changes are captured in source systems as they happen and the changes flow to the target systems. Attempts are made to keep all information up-to-date and in sync.*

- *Eliminating batch windows and enabling efficient data integration – IBM's unique, log-based change data capture (CDC) technology captures changed data directly from the database logs. This ensures that the performance of mission-critical applications running on the source system isn't adversely impacted and users can access these applications while integration processes are running. The integration process may be run throughout the day rather than being relegated to a nightly batch process. Change Data Capture technology also ensures that only changed data is replicated so users eliminate redundant data transfer, free up their batch window for other tasks, and save network bandwidth.*

© 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 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**