# Considerations for Building a Real-time Data Warehouse

By John Vandermay
Vice President of Development
DataMirror Corporation

**DataMirror** ®
The experience of now.™

# Considerations for Building a Real-time Data Warehouse

## John Vandermay, DataMirror Corporation

## Executive Summary

In today's fiercely competitive marketplace, companies have an insatiable need for information. Key to maintaining a competitive advantage is understanding what your customers want, what they need and the manner in which they want to receive your products or services. It is becoming increasingly clear that companies poised to experience the greatest success will be those firms that can effectively leverage their data to meet organizational needs, build solid relationships with stakeholders and above all, meet the demands of today's customers (Schroeck, 2000).

The global economy of today demands that organizations adhere to the constantly changing needs of the customer. Additionally, the speed and dynamic nature of business often negates the time required for long-term planning and time-consuming implementations in order to stay ahead. Because of this, organizations must implement solutions that can be deployed quickly and in a cost-effective manner (Zicker, 1998). So, how does an organization meet these ever-changing, complex requirements?

An effective real-time business intelligence infrastructure that leverages the power of a data warehouse can deliver value by helping companies enhance their customer experiences. Furthermore, a real-time data warehouse eliminates the data availability gap and enables organizations to concentrate on processing their valuable customer data. By designing a data warehouse with the end user in mind, you multiply your chances of better understanding what your customer needs and what you need to help that customer achieve his or her goals (Haisten, 2000).

The ability to access meaningful data in a timely, efficient manner through the use of familiar query and analysis tools is critical to realizing competitive advantages. Equally important is the moving and sharing of data throughout an organization, between departments, offices and business partners. But with the proliferation of mixed-system environments that must somehow be integrated with decision support systems, data marts and warehouses, electronic business solutions and enterprise applications, the challenges increase. When customer information is disjointed and spread across the organization, the challenges can become insurmountable.

Data – customer data, financial data, and Internet click-stream data – is a powerful asset provided it can be integrated and utilized to enhance customer experiences. Customers have become more complex and expectations are higher than ever before to meet their needs effectively. Today's businesses are under extreme pressure from both traditional and new rivals. Only those organizations that deliver the best customer experience will thrive. That means improving business management, providing market diversity and generating competitive advantages.

A successful real-time data warehouse can be the silver bullet your organization needs to prosper in the Internet era – that is, if you can avoid the common data warehousing pitfalls. In fact it has been said that data warehousing *is* e-Business. As we move forward, it is becoming clear that without the support of a

data warehouse, companies cannot successfully implement their e-Business strategies (Schroeck, 2000).

The following pages offer an informative approach to evaluating real-time replenishment software for feeding a data warehouse. We will outline real-time data transformation and integration requirements for the most functionally rich data warehouse and highlight how your business can experience positive results quickly to enable you to exceed the ever-changing needs and expectations of your customers.

## What is real-time data warehousing?

Until recently, there were few viable tools to provide real-time data warehousing nor an absolutely current picture of an organization's business and customer. But if businesses have survived without continuous, asynchronous, multi-point delivery of data in the past, why then would such solutions become so critical to business today?

Companies use data warehouses to store information for marketing, sales and manufacturing to help managers run the organization more effectively. The ability to manage and effectively present the volume of data tracked in today's business is the cornerstone of data warehousing. But when the data warehouse is replenished in real-time it empowers users by providing them with the most up-to-date information possible. Almost immediately after the original data is written, that data moves straight from the originating publisher to the data warehouse. Both the before and after image of a record is available in the data warehouse memory, thereby supporting easy and efficient processing for query and analysis at any time.

Given the benefits of real-time data warehousing, it is difficult to understand why the "snapshot" copy process has prevailed. Currently, the dominant method of replenishing data warehouses and data marts is to use extraction, transformation and load (ETL) tools that "pull" data from source systems periodically – at the end of a day, week, or month – and provide a "snapshot" of your business data at a given moment in time. That batch data is then loaded into a data warehouse table. During each cycle, the warehouse table is completely refreshed and the process is repeated no matter whether the data has changed or not.

Historically, best practices have been hampered by problems with integrating diverse production systems with the data warehouse. Snapshot copy was deemed "right" because it was next to impossible to get real-time, continuous data warehouse feeds from production systems. As well, since query tools were relatively unsophisticated and complex to debug, it was also difficult to get consistent, reliable results from query analyses if warehouse data was constantly changing.

In the Internet era, more people are beginning to realize the limitations that snapshot copy replenishment presents and demand better alternatives. Snapshots do not involve entire database movement but simple captures of parts of database tables; for example, specified columns. As well, not each individual change is made to a record between copy processes. In this light, the snapshot process can be likened to looking at last week's newspaper or using last week's stock market results to trade stock today.

The Internet era is about having absolutely current and up-to-date business intelligence information. Data is a perishable commodity: the older it is, the less relevant. Businesses need tools that can provide real-time business intelligence and an absolutely current and comprehensive picture of their organization and their customers – not last week or last month, but right now.

# Components of Real-Time Data Warehousing

An up-to-the-second view of customer data, once an ideal, is fast becoming a reality for businesses wishing to implement real-time business intelligence solutions. But how does the data warehouse actually operate?

An intelligent warehousing solution and framework can commonly be divided into three fundamental tiers with data flows between them. The three layers are Presentation Layer, Architecture Layer, and Middleware Layer. These tiers or layers must be seamlessly integrated and function as one to ensure the immediate success and long-term benefits of a data warehouse.

## Presentation Layer

The presentation layer manages the flow of information from the warehouse to the analyst, providing an interface that makes it easier for the analyst to view and work with the data.

This layer is where graphical user interface (GUI) tools are most important. Front-end query tools should provide an easy and efficient way to visually represent data for decision making in two or more dimensions. Pattern recognition and analytic algorithms can highlight areas for close human analysis, but in the end humans still have an edge in improvisation, gut feeling and trend forecasting. Warehousing assists users in the analysis of sales data so they can make informed decisions that have real-time impact on company performance.

The presentation layer's ability to store and present multidimensional views or summaries of data is one reason why multidimensional databases and query tools are popular at this level of the warehouse.

## Architecture Layer (Structure, Content/Meaning)

The architecture layer describes the structure of the data in the warehouse. An important component of the architecture layer is flexibility. The level of flexibility is measured in terms of how easy it is for the analyst to break out of the standard representation of information offered by the warehouse in order to do custom analysis. Custom analysis is where semantic thickness becomes important.

Semantic thickness is the degree of clear business meaning embedded in both the database structure and the content of the data itself. Field names such as "F001" for customer number and obscure numbers such as "01" to indicate "Backorder" status are considered semantically thin, or ambiguous and difficult to understand. In contrast, field naming standards such as "Customer_Name" containing the full customer name and "Order_Status" containing the complete description "Backorder&" are semantically thick, meaningful and easily understood.

In other words, data structure and content must be clear to the analyst at the presentation layer of the data warehouse. The underlying data schema for the warehouse should be simple and easily understood by the end user of the data.

## Middleware Layer (Interfaces and Replenishment)

The middleware layer is the glue that holds the data warehouse together. It integrates the data warehouse with production and operational systems. Data needed for warehouse applications often must be copied to and from computers of different types in different locations. Warehousing often implies transformational data integration. Production data needs to be secure and is frequently not in the format needed for warehousing. Real-time integration and replenishment tools that help businesses deal with the data management issues of implementing a data warehouse can add real value.

The rest of this paper will focus on how a real-time integration and replenishment solution – or a capture, transform and flow (CTF) tool can contribute to the simplicity and efficiency of a real-time data warehouse.

# Efficient Business Intelligence Replenishment

Before the Internet existed, only a few dozen users – usually hardcore data analysts – accessed most data warehouses. But the democratization of information access over the last few years has created new challenges. Web-based architectures must routinely handle large volumes of concurrent requests while maintaining consistent query response times, and must scale seamlessly as the data volume and number of users grows over time. In addition, data warehouses need to remain available 24 hours a day because the web makes it cost effective for global corporations to provide data access capabilities to end users located anywhere in the world. This is where data replenishment and resiliency tools come in to provide real-time access.

Even if your organization does not wish to implement real-time data warehousing, it is still important to consider how efficient your current extract tool is at replenishing the data warehouse. Most ETL tools are batch processors, not real-time engines. This method can be both resource intensive and time consuming. The larger the data warehouse, the longer it takes to replenish with this method. In some cases, the volume of data being loaded into the warehouse begins to exceed the batch window allotted for it. This process is inherently inefficient. In a typical environment where 20 per cent of the data on productions systems is changing every week or month, why would you choose to refresh the entire data warehouse every week or month? Why send 20 gig of data when you could send only 20 per cent of 20 gig – or 4 gig – that has changed since the data warehouse was last replenished?

Due to the growth in business and the related increases in data, companies may find it difficult to fit its "batch job" into a periodic time window of eight hours, and may as a result cut into normal usage hours. In contrast to standard ETL tools, consider an advanced CTF solution that instead captures, transforms and flows data in real-time into an efficient, continuously replenished data warehouse.
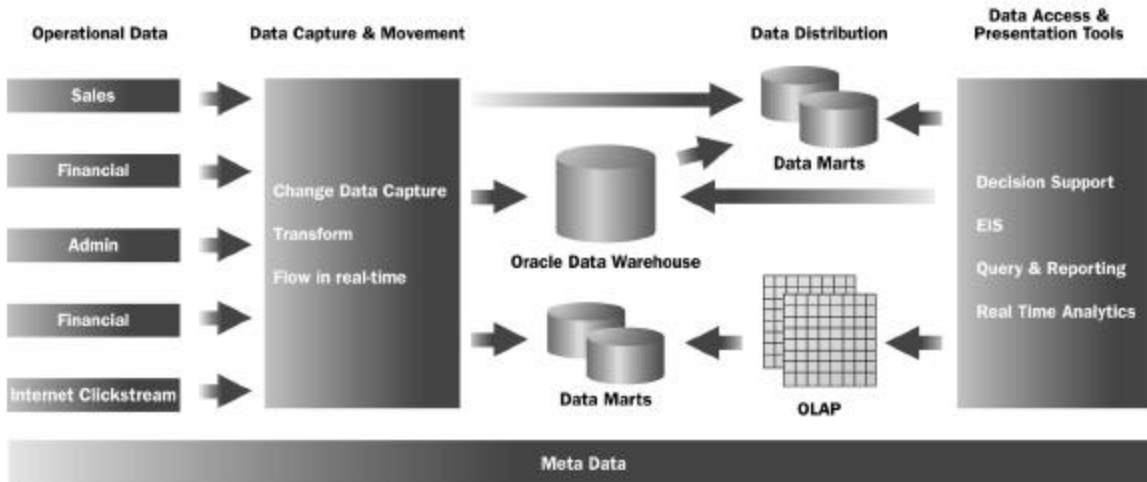
*Figure 1: Typical data warehouse implementation utilizing
capture, transform and flow (CTF) technology for real-time replenishment.*

# Capture, Transform and Flow (CTF)

## Change Data Capture

Today, more and more businesses using a data warehouse are beginning to realize they cannot achieve point-in-time consistency without continuous, real-time change data capture. There are several techniques used by data integration / replenishment software to move data. Essentially, integration tools either push or pull data on an event driven or polling basis.

Push integration is initiated at the source for each subscribed target. This means that as changes occur, they are captured and sent, or "pushed" across to each target. Pull integration is initiated at the target by each subscribed target. In other words, the target system extracts the captured changes and "pulls" them down to the local database. Push integration is more efficient as it can better manage system resources. As the number of targets increases, pull integration becomes resource draining on the source system, especially if that machine is a production machine that may already be overworked.

Event driven integration is a technique that involves events at the source initiating capture and transmission of changes. Polling involves a monitoring process that polls the status to initiate capture and application of database changes. Event driven integration conserves system resources as integration only occurs after preset events whereas polling requires continuous resource utilization by a monitoring utility.

But in order to compete with an information-driven Internet era, organizations must employ solutions that offer the option of updating databases as incremental changes occur, reflecting those changes to subscribed systems. With advanced CTF solutions, every time an add, change or delete occurs in the production environment, it is automatically captured and integrated or "pushed" in real-time to the data warehouse. By significantly reducing batch window requirements and instead making incremental updates, users regain computing time once lost.

Beyond real-time integration, change data capture can also be done periodically. Data can be captured and then stored until a predetermined integration time. For example, an organization may schedule its

refreshes of full tables or changes to tables to be integrated hourly or nightly. Only data that has changed since the previous integration needs to be transformed and transported to the subscriber. The data warehouse can therefore be kept current and consistent with the source databases.

## Transformation

The way companies think about data, and the way it is represented in databases, has changed significantly over time. Obscure naming conventions, dissimilar coding for the same item (e.g. number representation as well as character based codes), and separate architectures are all commonplace. Software that can transform data across multiple computing environments and databases can remedy these problems while consolidating the information in your data warehouse.

Companies are beginning to realize the benefits of sharing data between enterprise resource planning (ERP) systems and relational data stores housed in databases including Oracle, Sybase, DB2/UDB, and Microsoft SQL Server. The problem is that ERP systems use proprietary data structures that need to be cleansed and reformatted to fit conventional database architectures. Rows and columns may have to be split or merged depending on the database format. For example, an ERP system may require that "Zip Code" and "State" is part of the same column while your company's data structure may have the two columns separated. Similarly, your company may have "Product Type" and "Model Number" in an inventory database as one column, whereas the ERP system requires them to be split. Data transformation and integration software can accommodate these requirements in order to make your data – and consequently, your data warehouse – more useful and meaningful to users.

| Publisher | Transformation process | Subscriber |
|---|---|---|
| Smith, M.<br>$10.00 U.S.<br>20"<br>1B | Two-field Consolidation and Rearrangement<br>Euro Conversion<br>Unit Conversion<br>Value Substitution | Mary Smith<br>9.333 Euros<br>50.8 cm<br>In Stock |

*Figure 2: Sample data transformations. Data warehousing projects often require operational data to be reformatted, enhanced and standardized in order to optimize warehouse performance and make business intelligence content more meaningful to end users. Effective real-time replenishment tools should perform data transformations on-the-fly without requiring data staging.*

Other applications of data transformation software include changing data representation (U.S. dollars converted to British Sterling, metric to standard, character columns to numeric, abbreviations to full text, number codes to text), visualization (aggregate, consolidate, summarize data values), and preparation for loading multidimensional databases. Transformational data integration software can conduct individual tasks such as translating values, deriving new calculated fields, joining tables at source, converting date fields, and reformatting field sizes, table names and data types. All of these functions allow for code conversion, removal of ambiguity and confusion associated with data, standardization, measurement conversions, and consolidating dissimilar data structures for data consistency.

## Flow

This refers to replenishing the feed of transformed data in real-time from multiple operational systems to one or more subscriber systems. Whether a data warehouse or several data marts, the flow process is a smooth, continuous stream of bits of information as opposed to the batch loading of data performed by ETL tools.

# Evaluation Criteria for an Efficient Real-Time CTF Solution

Real-time capabilities are ideal in maintaining a current record of a customer profile and customer needs. But many vendors offer tools that are capable only of straight database copies, unidirectional integration, or database snapshots. While these tools are useful for some projects, they often fall short once the organization begins to outgrow its original transformation needs.

More robust and powerful capture, transform and flow (CTF) software exists that can facilitate the real-time delivery of meaningful information to subscribed systems, movement among heterogeneous platforms and databases, and the selecting and filtering of the data transmitted. In the ever-changing Internet era, companies should strive to select real-time CTF solutions that offer the following capabilities and features.

## Selectivity

Business solutions like data marts and warehouses require the ability to select and filter which data is moved throughout the organization. Efficient CTF software offers an array of features including built-in data filtering and selection functions in addition to data enhancement and transformation capabilities. This allows source data to be selectively filtered by row and/or column before replenishing the data warehouse.

For example, a company planning to implement a data warehouse may want to filter out information about salary, sick time or vacation, while allowing access to hire date, salary range, and department. By limiting access to sensitive information, row and column selection enables users to populate data warehouses and data marts with user-centric information or integrate location-specific data to particular sites.

## Support for Heterogeneous Environments

Because of changes in technology, corporate mergers, and acquisitions, most organizations now bear multiple computing platforms and databases, each storing separate pockets of information that may be entirely incompatible to the next.

Most data integration solutions focus on moving data only between homogeneous systems and database software. Some integration tools, however, are capable of moving data among a wide range of databases and systems including Oracle, DB2/UDB, Sybase, and Microsoft SQL Server across OS/390, OS/400, UNIX, Linux and Microsoft Windows NT/2000. A number of vendors currently offer transformational data integration tools to consolidate and synchronize heterogeneous data into a data mart or warehouse, or to integrate data from legacy production systems with data from newer web-based systems. However, many of these tools do not offer real-time capabilities.

## Ease of Administration

Some integration software can be awkward to install and set up. A serious consideration when evaluating transformational data integration tools is the work required for setup and implementation of the software. IT staff and end users seek an "out-of-the-box" experience from data warehousing tools. Organizations should ensure that no programming changes to existing applications and databases are required. A good thing to keep in mind is that the solution software is meant to avoid time-consuming, resource-intensive, and costly custom extract programming. Vendors that offer portable graphical user interfaces and utilities for web or wireless management of integration software can empower businesses and ease the strain of constantly being on-site to view and control integration status. Organizations can therefore focus on driving their business–not their technology.

## Meta-data Management Capabilities

Meta-data is information about data. It allows business users as well as technical administrators to track the lineage of the data they are using. Meta-data provides information about where the data came from, when it was delivered, what happened to it during transport, and other descriptions can all be tracked. What can you do with meta-data? It is widely held that there are two types of meta-data: technical, or administrative meta-data, and business meta-data. Administrative meta-data includes information about such things as data source, update times and any extraction rules and cleansing routines performed on the data. Business meta-data, on the other hand, allows users to get a more clear understanding of the data on which their decisions are based. Information about calculations performed on the data date and time stamps as well as meta-data about the graphic elements of data analysis generated by front end query tools. Both types of meta-data are essential to a successful data mart or warehouse solution. How well the data warehouse replenishment solution you choose manages and integrates meta-data may affect the performance of presentation tools and the overall effectiveness of the data warehouse.

Essentially, examining meta-data enhances the end user's understanding of the data they are using. It can also facilitate valuable "what if" analysis on the impact of changing data schemas and other elements. For administrators, meta-data helps them ensure data accuracy, integrity and consistency. During data transport, advanced capture, transform and flow (CTF) replenishment solutions store meta-data in tables located in the publisher and/or subscriber database. This is an attractive feature to companies wanting to share meta-data among heterogeneous applications and databases. Most tools and databases manage meta-data differently. That is, they store the meta-data in distinct formats and use proprietary meta-data to perform certain tasks. An open CTF solution allows organizations to distribute meta-data in different formats using published industry standards. Using CTF technology based on open standards such as the Meta-data Coalition Standard for Meta-data Integration, Open information Model (OIM), or XML Interchange Format (XIF), meta-data can be easily integrated to another repository in the required format. This addresses the challenge of standardizing meta-data. Without this functionality, companies often must resort to custom development of a tool capable of entering meta-data in a variety of formats – a development task that often proves time-consuming, difficult to accomplish, and may negatively impact time to market for the business intelligence application.

## Scalability

Scalability refers to the ability of a computer system or a database to operate efficiently with larger quantities of data. While it is possible to predict some of the future information needs of an organization, a large portion will be unpredictable. In the Internet era, it is certain that as business environments change, so too will the kinds of decisions that need to be made and the information that influences them. Therefore, businesses need to think ahead and look for a scalable solution when searching for an efficient data integration tool. Organizations should not look at data warehouse development as having a beginning, middle and end, but as a continuous process that must evolve with the organization. Failing to do so will result in the loss of information necessary for future strategic decision making and competitive advantage.

# Current Issues in Real-time Data Warehousing

## Time to Market

In today's competitive economy, time to market means everything for data warehousing projects. Why do so many data warehousing projects fall behind schedule or even fail? One major reason for data warehouse failures is that many data warehouses are populated with operational data that is poor in quality. Raw operational data needs to be selected, filtered and transformed before consolidating it in warehouse tables for business intelligence purposes. Operational data is typically stored in multiple tables and consists of codes and abbreviations, making it difficult to access for decision support. A simple invoice, for example, may contain data from over a dozen different tables. More often than not, operational systems also contain inconsistent data. An inventory system may store data as "Male" and "Female," while a system used by sales stores the same information as "M" and "F."

Given the circumstances, most would agree that unleashing end users on a data warehouse without first cleansing or transforming the raw transactional data that populates the warehouse would be a poor idea. Data quality can also seriously affect data warehouse performance. With data warehouses and data marts, the analogy is: "garbage in, garbage out." You won't find the trends and relationships you're looking for in the data unless you feed your query and analysis tools with the right information. A little knowledge, or the wrong knowledge, can be a dangerous thing. It can give you an incomplete or flawed picture of your customer or your business.

This problem of data quality can be avoided by selecting a replenishment solution that offers advanced capture, transform and flow (CTF) technology. CTF tools enable you to capture raw data from multiple operational databases and flow the data in real-time into data warehouse tables while transforming the data on-the-fly into meaningful information. Users are empowered through the means to translate values, derive new calculated fields, reformat field sizes, table names and data types. CTF tools help accelerate time to market while adding value to business intelligence information by keeping the data clean, current and in a format conducive to query and analysis. Through a combination of best practices and best-of-breed solutions such as capture, transform and flow tools, companies can reasonably expect to have end users querying the data warehouse within a short time frame.

In addition, data warehouse projects can quickly get out of hand in terms of size and scope because it is difficult for IT managers to deny requests from users and executives to change or expand the design. This can make it difficult or even impossible to deliver BI systems on time and under budget. Large data warehouses need to integrate data from across the organization—different hardware, operating systems, databases and applications—and integration efforts can be time-consuming and costly. A successful BI application takes good advance planning, organization-wide sponsorship and input, and doing your homework to choose the right tools and the right vendors to partner with.

The remainder of this white paper will explore other issues that should be considered when the goal is to implement a data warehouse project on schedule and on budget.

## Buy versus Build

Despite the buy-versus-make recommendations of most data warehousing analysts, some companies still choose custom coding to handle the replenishment and transformation of data. In this situation, businesses can write customized programs to integrate data to the data warehouse. However, given the productivity gains of using a mapping-based integration tool over scripting, it is difficult to justify allocating the time and resources required for developing custom applications. Mapping-based tools involve built-in administration utilities that provide quick and easy definition of the integration process. Changes in the integration schema are simply re-mapped using the same front-end that defined the original process.

Some companies who choose to bring the task in-house and custom code extraction routines soon discover that it is a difficult problem with its own challenges. Why allocate a programmer to do punishing extract programming when your organization can use an easy-to-implement mapping-based tool? Packaged solutions can provide businesses with many advantages including flexibility, scalability, and upgrade support for the latest database versions upon their release.

## Aggregation

Aggregation refers to the gathering of information in separate sets from two or more sources. Often, this data is stored in a data warehouse in a summarized form. For example, an organization may wish to summarize the data by various time periods. Aggregates are used for two primary reasons. One is to save storage space; data warehouses can get very large, and the use of aggregates greatly reduces the space needed to store data. The second reason is to improve the performance of business intelligence tools. When queries run faster, they take up less processing time and the users get their information back much more quickly.

However, there is also a negative side to data aggregation. The process can result in the loss of time-sensitive linear data. If a business is trying to compile a complete customer profile in order to understand its customers, aggregation is not beneficial.

Some data warehouses store both detailed information and aggregated information. This approach may take up even more space, but gives users the possibility of looking at all details while still having good query performance when looking at summaries. As well, user exit support exists today in some software that facilitates aggregation while also providing the ability to create a complete customer profile.

# Summary

When deciding how to build a data warehouse, there is a tendency to get caught up in the technology and forget the basics that contribute to its creation. The questions that your business needs to ask are how fast does your company respond to new ideas, new competitive pressures and new opportunities? How can it do great work fast? Can each branch of the organization stay informed about customers, competitors and business trends?

The competitive and customer pressures of the new economy have created an insatiable demand for up-to-the-second business information. It is no longer acceptable for businesses to make decisions based on day-old or week-old data. Employees, decision-makers, partners and customers alike need access to information while it is still relevant. Real-time data warehousing essentially allows organizations to combine long-term planning tactics with up-to-the minute decision-making to ensure enhanced customer experiences, influence customer loyalty and increase the bottom line (Brobst and Venkatesa, 1999).

The Internet has significantly changed the scope and expectations for data warehousing implementations. Developing a robust, scalable data warehouse with consistent, real-time data and the ability to answer all user-information needs is not an easy goal to achieve. Keeping data current is one of the most difficult and time-consuming challenges in managing data warehouses and data marts. But if organizations fail to take advantage of real-time data capabilities for business intelligence, they will lose the opportunity to respond quickly to changing market trends. These businesses will be less agile and will have difficulty meeting the customers' rising expectations for 24/7 service. Predictably, they will lose market share to truly customer-centric businesses that do invest in real-time data warehousing.

By understanding the relevant issues and staying informed of current practices in data warehouse development, your organization can harness the power of real-time business intelligence and attain a cost-efficient, profitable computing environment. In a very short time, your business could be well on its way toward meeting customer needs and generating a competitive edge in today's competitive economy.

# References

Brobst, Stephen A. and Venkatesa, AVR, "What is Active Data Warehousing?," Abridged from "Active Warehousing," Teradata Review Spring 1999 <http://www.ncr.com/repository/articles/data_warehousing/what_is_active_dw.htm>.


Haisten, Michael, "Real-Time Data Warehouse: What is Real Time about Real-Time Data Warehousing?," DM Review Online August 2000 <http://www.dmreview.com/master_sponsor.cfm?NavID=68&EdID=2589>.


Schroeck, Michael J., "E-Analytics—The Next Generation of Data Warehousing," DM Review August 2000 <http://www.dmreview.com/master.cfm?NavID=55&EdID=2551>.


Zicker, John, "Real-Time Data Warehousing," DM Review March 1998 <http://www.dmreview.com/master_sponsor.cfm?NavID=55&EdID=676>.

## About the Author

John Vandermay is Vice President of Development for DataMirror Corporation where he is responsible for all software development activities across the DataMirror product family as well as quality assurance. Prior to joining DataMirror, Mr. Vandermay held senior development and management positions at Marcam Solutions including Director of Development. As Director, Mr. Vandermay was responsible for development in Toronto in conjunction with other development sites in Boston and Tokyo and liaised with many large customers including Nestle, NEC and 3M. Before joining Marcam, Mr. Vandermay worked for four years at Clarica (formerly The Mutual Group) in Waterloo, Canada developing mainframe applications and Windows client-server programs. Mr. Vandermay graduated from Wilfrid Laurier University in 1989 with a Bachelor of Science degree in Physics and Computer Science. Mr Vandermay may be reached by phone at 905-415-0310 or by e-mail at jvandermay@datamirror.com.

# About DataMirror Corporation

DataMirror (Nasdaq: DMCX; TSE: DMC) delivers solutions that let customers integrate their data across their enterprises. DataMirror's comprehensive family of products unlocks *The experience of now ™* by providing advanced real-time capture, transform and flow (CTF) technology that gives customers the instant access, integration and availability they demand today across all computers in their business.

Over 1,500 companies use DataMirror to integrate their data. Real-time data drives all business. DataMirror is headquartered in Toronto, Canada, and has offices worldwide. DataMirror has been ranked in the Deloitte and Touche Fast 500 as one of the fastest growing technology companies in North America.

**r** DATAMIRROR RESOURCE CENTER
SOFTWARE I SOLUTIONS I BEST PRACTICES
www.datamirror.com/resourcecenter

FOR MORE INFORMATION CALL 1 800 362-5955