

# Look before you leap

Activating the data warehouse requires careful consideration. *by Roger Mann*

The industry is abuzz over active data warehousing, a process also known as “active enterprise intelligence,” “pervasive business intelligence (BI)” and “operational BI.” As your organization moves into this arena, a number of critical success factors need to be considered.

Active data warehousing is an evolution of the enterprise data warehouse (EDW). The traditional EDW collects and loads volumes of transactional data nightly from across the organization. This data is then segmented, categorized and made available for analysis via reports and other tools. (See figure 1.) The organization’s decision makers, normally back-office users, analyze this historical data to help determine the strategic direction of the enterprise.

Business decisions made by front-line operational users like bank tellers or call center employees, on the other hand, require up-to-date transactional information as well as historical data derived from the data warehouse with response times of seconds or sub-seconds. In the past, this demand has stuck IT with a dilemma: The online transaction processing (OLTP) systems did *not* have the history but *could* meet the response-time requirements; a traditional EDW *had* the history but could *not* meet the response times.

To satisfy the history and response-time requirements, IT professionals created operational data stores (ODS). These ODS databases require the duplication of both transactional and data warehouse data into yet another database. But this results in

higher development and maintenance costs and a database that can often deliver inconsistent information.

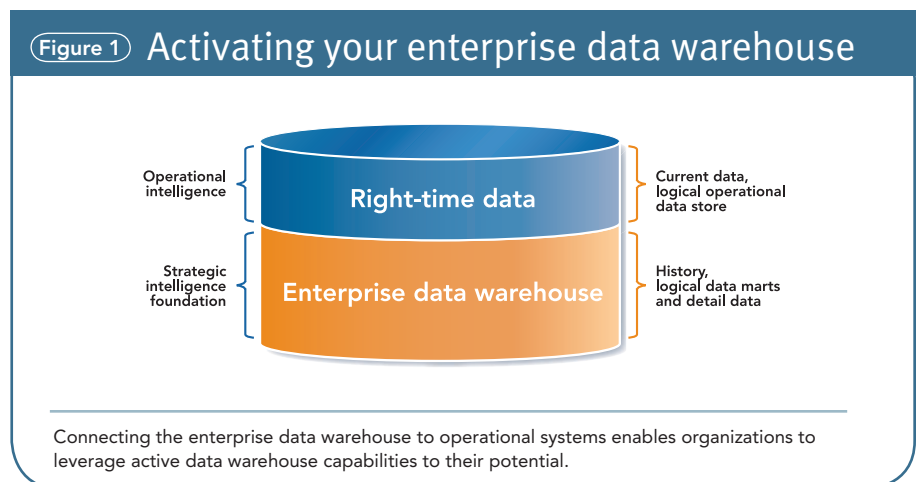
Teradata considered this dilemma and engineered specific features and functionalities into its relational database engine that allow companies to consolidate the EDW and ODS into one database management system and meet the response-time requirements. This process activates the data warehouse, minimizes and often eliminates costly data duplication and maintenance, and provides a single view of the business to strategic (analytical) and tactical (front-line operational) users. The result is data that is timelier, more consistent, more accurate and less costly to the enterprise.

## 6 components of activation

Transforming a data warehouse into one that is active involves six components (see figure 2, page 69), each of which differs

significantly from its counterpart in a traditional data warehouse:

- 1. Active access.** Data warehouse access is given to hundreds or possibly thousands of operational users via tactical applications that require response times measured in seconds or sub-seconds.
- 2. Active loading.** Data loading needs to be as near real time as is required by the particular application. For some, this entails loading data every hour, while other applications require the data to be loaded within minutes, seconds or even sub-seconds of the OLTP systems. The classic extract, transform and load (ETL) process must be modified to accommodate near real-time integration of data from multiple sources.
- 3. Active enterprise integration.** The data warehouse must integrate with



portals, Web services, Web sites and other operational environments, such as enterprise application integration tools.

**4. Active events.** The most potentially explosive frontier in the active data warehouse portfolio is processing active events. As the data warehouse ingests the volumes of transactions each day, this component can identify exceptions that warrant immediate action. By building triggers and stored procedures into the loading process, the data warehouse can identify these situations and alert the enterprise to take action.

**5. Active workload management.** The integration of strategic/analytical processing with tactical processing expands the breadth needed for workload management strategy. The active data warehouse has service level agreements (SLAs) that require data to be loaded in minutes and response times in seconds or sub-seconds while the strategic EDW applications still need to meet their expected SLAs.

**6. Active availability.** As the data warehouse moves out of the back office and into the operational arena it will immediately become a system in which high-availability and disaster recovery methodologies become critical. The system must have the highest levels of availability with clearly defined backup, archive and restore disaster recovery tools and failover functionality. Unfortunately, disaster recovery is often overlooked until a disaster underscores its absence.

Teradata originated as a platform to drive analytical decisions by providing massive parallelism to explore historical business data. With its long history of incorporating specific functions and features into the relational database engine, Teradata has seen the value of enabling

### Going active

A Web-based company processed about 300,000-plus queries per day with its Teradata system. After adding an "active" application that accessed customer relationship management data and required sub-second response time, the system's workload immediately increased. Thousands of new users were introduced to the system and the average number of queries added to the mix was more than 1.3 million per day. Yet even with that many daily queries, the new application consumed less than 4% of the total system resources.

this activating capability. These functions and features include:

- > Advanced indexing strategies and support
- > Inherent workload management tools
- > Near real-time data load tools, strategies and support
- > Tactical query strategies and support

An organization that has initiated one or more of these components is at a distinct competitive advantage as it will be able to deliver consistent, accurate and timelier information to its strategic and tactical users, and be able to address its service level requirements.

### Cultural changes are necessary

Activating the data warehouse environment by adding operational application capabilities transforms it from a back-office reporting system to an operational system with increased visibility and demand. Front-line users require up-to-date data, near real-time performance and potential availability of 24 hours a day, every day of the year. Receiving timely

and accurate data will allow users to reach the goal of faster, smarter decision making, but getting to that point will require cultural changes within the organization.

The success of implementing this activation process will depend on how easily an organization can transform its culture. An active data warehouse requires most of the infrastructure of an operational environment, as well as governance, to ensure its viability. The enterprise's entire personnel, from the organization's top-line management to the back-office and front-line workers, must be educated on the importance and purpose of the EDW and be on board with the rules and requirements established by the company's governance team.

The key player in this equation is the CIO. Skeptics in the organization who have vested interests in legacy architectures will surface, so CIO support is necessary to overcome these naysayers and propel the traditional data warehouse into an active environment.

### Educate the organization

No one would consider beginning a project blindly. This is true of building the EDW and extending its functionality to include front-line tactical users. The intricacies and considerations of creating a world-class data warehouse capable of supporting tactical workloads require attention to details not previously considered.

Every organization is different based on its workload, data, business rules and so on, but all of them can discover through a number of sources helpful information on how to establish an active data warehouse environment. Networking with peers from other corporations is beneficial for sharing ideas and processes. Attending user groups and industry event forums can provide information on what has or has not worked for others—but be sure to understand the reasons behind

this information, as the condition may not apply to your environment.

Developing in-house education that addresses business and IT needs and providing it to the appropriate users is critical. Numerous online classes are available through Teradata Support Services to help educate your organization on how to activate your data warehouse. A series of online presentations details how the six activation components interact in your environment.

Business and IT employees must fully understand what data is available, what they can expect from the system and the benefits an active data warehouse can provide.

## Project selection

Evolving the data warehouse from a traditional decision support system to one that extends to front-line users usually happens somewhat naturally. When new applications are being developed, it makes sense to consider the EDW, as it will need to accommodate the growing spectrum of business data. While most of us in data warehousing see this

development as a good thing, depending on the degree of discipline surrounding the existing data warehouse, exercising some caution is advised.

The first foray into active data warehousing should be with a project of minimal risk. This initial venture will enable you to learn more about your active data warehouse environment and identify the current processes that will continue to work, as well as those that need adjusting. From this exercise, you can also determine what processes are missing and need to be developed. While it won't help you understand and identify everything you need for future projects, this first effort will provide a great deal of education and revelation and will put you in a better position for success in the next project.

Labeling this first project as a proof of concept (POC) is one way to extract the greatest benefit. As a POC, it becomes a learning experience in which everyone participates, and the pressure surrounding the project is minimized. Business users

and IT personnel will come together with a common goal of exploiting the data warehouse and developing the application, and the ensuing open conversations will provide a crucial building block for an ongoing partnership between business and IT.

## Best practice

Typically, the EDW mantra is to answer any question on any data at any time. Moving toward an active environment will introduce into

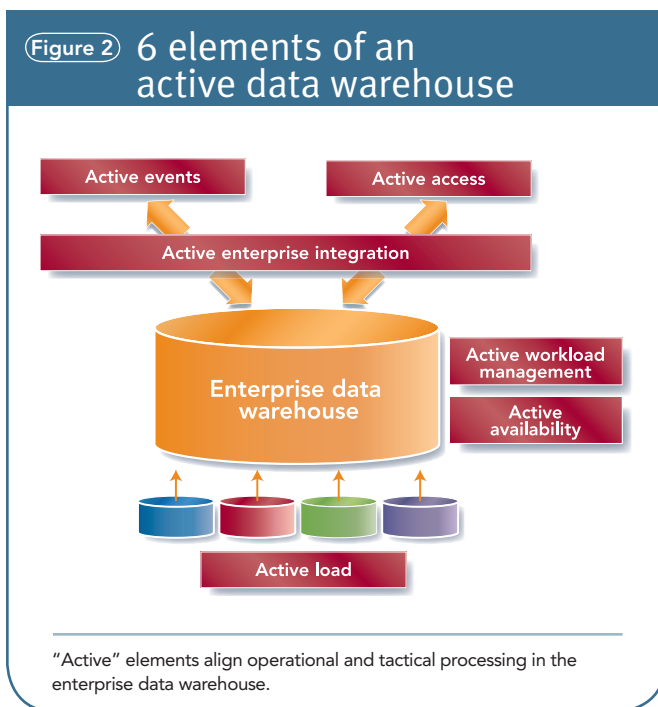
that workload a set of targeted queries that require response times of seconds or sub-seconds. As the number of users executing these queries grows, and the number and speed of queries escalates, the workload volumes can shock corporations. What must be remembered is that these are very targeted, highly tuned queries requiring very few resources, so the impact of these queries on the data warehouse tends to be minimal.

If rigorous program testing, change control, performance monitoring, problem detection and escalation, ongoing environment monitoring and disaster recovery are not part of the existing data warehouse environment, they must be addressed for the new active environment. Activating the data warehouse takes it from back-office to front-office processing. Indeed, active applications move the data warehouse to mission-critical status needing 24x7 availability and will require careful examination, as will the organization's business needs, to determine the best way to safeguard the business and ensure availability.

## Ready, set, activate!

Moving to an active data warehouse environment is a new and exciting frontier for IT specialists. Long-term advantages for the entire organization can be realized when operational systems and traditional data warehousing consolidate, and the architecture that holds the single view of the business is leveraged for both strategic and operational applications. **T**

*Roger Mann started with Teradata in 1983 and works in the Active Enterprise Integration Center of Expertise.*



**Online** To learn more about activating the data warehouse, go to Support Services on [Teradata.com](http://Teradata.com).