

Know sooner rather than later

Low-latency data paves the way for better decision making. *by Richard Hackathorn and Jack Garzella*

Business happens quickly in the global economy, and the pace is increasing. A critical issue facing every organization is the freshness of its business data.

Recent studies have shown a significant business requirement for data “freshness,” specifically, data that is less than 24 hours old. Technically, the speed in which data is captured is referred to as data latency, and is defined as the time interval from a business event (such as an order placed by customer) to the business response (such as shipping the order to the customer).

Businesses typically refresh their data on a daily cycle by acquiring it from transaction systems and feeding it into their enterprise data warehouse (EDW) for analysis. Data refreshed more often than daily is labeled “low latency” and may lag business events from several seconds to several hours.

Measuring data latency in minutes is increasingly required to support critical operational processes—the heart of business intelligence (BI). Knowing about your business sooner, rather than later, provides a competitive edge and enables you to better manage ordinary as well as unexpected situations.

Using data to stay ahead

There is business value when we can do something different within our enterprise today rather than wait until tomorrow. Organizations generally have the necessary

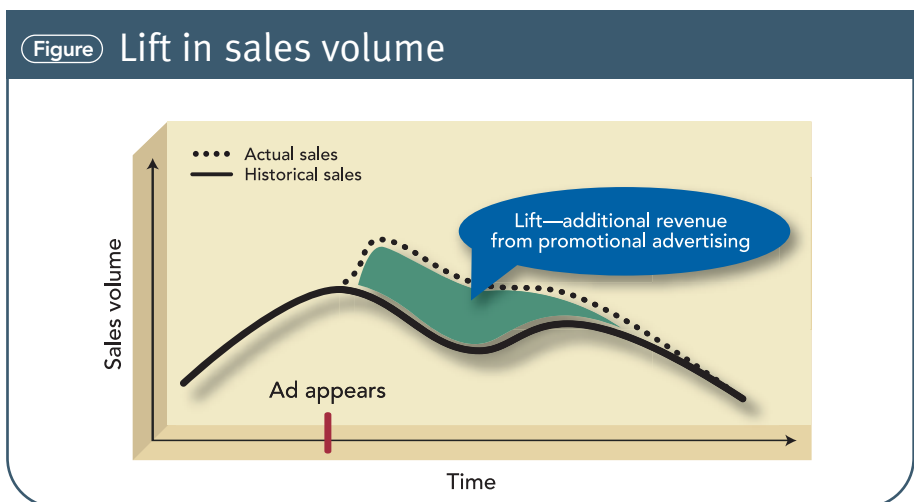
data to efficiently run the business. However, to be competitive, the key questions are: “What do we do with what we know?” “What can we do differently to improve our business if we have the proper information available immediately?”

Let’s consider an example in which we portray a large retailer: In an attempt to increase our sales, we release a one- to two-day special promotion over the holidays. To validate the effectiveness of the promotion, we will monitor the sales lift of the campaign by analyzing the changing data. (See figure, below.)

If this is a traditional campaign in which material is mailed to prospective

customers, our monitoring timeline might be a week. In this case, we would likely use the analysis to adjust the next campaign scheduled to be launched the following month for a similar item.

In a second scenario, if this is an e-mail or Web campaign (in which a promotion window pops up on the home page), our monitoring timeline is just a few hours. Again, we use the available data to decide how to improve our campaign. In either case, the question is: “What can we do with the information provided about the sales volume?” The answer depends on when we obtain this analysis.



The solid line shows expected sales based on historical data. By providing low-latency data, businesses can make timely decisions on how to manage a marketing campaign so that a “lift” of additional revenue will replace the expected dip in sales.

More options with fresh data

For traditional BI, the data warehouse would be loaded nightly with the data from that day's sales and we would receive the analysis the following morning. Consequently, in an attempt to increase the lift, we'd probably adjust the pop-up promotions for the next day. Note that there is nothing we can do about yesterday's campaign—it's history!

Ideally, however, we would get the analysis as soon as technically possible. In this best-case scenario, sales updates would continuously be loaded into the data warehouse, making analysis available within minutes of the sale event. We could watch the analysis as a dashboard display that is refreshed minute by minute. Within an hour, we could compare this promotional campaign against similar ones.

We now have more options. If the campaign is successful, we could continue it without making changes. If the campaign is too successful with an unusually high volume, we may have set the price too low. An immediate price increase may boost overall profitability with similar sales volume. Conversely, if the campaign is faltering with a volume that is too low, we could decrease the price, change the messaging or offer a perk, such as free shipping.

Low latency, better decisions

This walk-through illustrates how low-latency data not only speeds up the execution of a business process but changes the business value altogether. With data loaded once a day, a two-day campaign can be adjusted only once. With continuously fresh data, the retailer can adjust constantly for the maximum return.

Business information processing involves three steps. First, we capture the data and acquire it into the data warehouse for analysis—capture latency. Second, we analyze the data—analysis latency. Third, we make a decision and take an action—decision latency.

Incorporating low-latency data into our business processes increases the opportunity for more frequent data analysis, thereby making more decision options available. When we, as organization leaders, are willing and able to manage based on a new set of standards and enable our staff with new skills, we can realize the full potential of low-latency data. **T**

Richard Hackathorn is president and founder of Bolder Technology, a consulting and

education firm specializing in business intelligence and data warehousing. As an author, Richard writes for the Business Intelligence Network and has published three books. He presents regularly at leading trade conferences and has conducted professional seminars in eighteen countries.

Jack Garzella is a managing partner for Teradata in the e-business and telecommunication vertical and has managed the building of more than 15 data warehouses.

Realizing the benefits of low-latency data

Getting organizations to realize the benefits of low-latency data is essential to businesses' ongoing success. There is widespread belief that near real-time data is too expensive for most business applications. This belief, however, is no longer valid in many situations because of the rapid evolution of data acquisition technology. A recent study conducted by Bolder Technology and sponsored by GoldenGate Software analyzed the infrastructure cost for supporting low-latency data in an enterprise data warehouse (EDW). The objective of this study was to educate IT professionals to ask the right questions and for business executives to understand what is technically possible.

The study explored five typical architectures for data acquisition and estimated the cost factors for each. Subjective estimates for costs were used based on firsthand experiences with several large systems in production today.

The study concluded that it is often more cost-effective to use data acquired through continuous stream technology—as in an EDW—than it is to use data acquired through traditional technology. In fact, the results pointed to the similarity in cost between continuous stream data acquisition and upgrading current extract, transform and load technology to mini-batch processing. Not only does an EDW offer a simpler architecture, but it also provides greater ability to support future requirements.

Rather than form general conclusions based on these estimates, this study suggests a framework and method for evaluating the trade-offs in real situations. The full report for the study can be found at www.b-eye-network.co.uk/view-whitepapers/3981 along with the spreadsheet template used to compile the cost estimates.

Technology advances are changing the data acquisition architectures for enterprise data warehousing. The accepted industry wisdom about low-latency data is rapidly shifting. Organizations are finding that an architecture that supports continuous stream updates with near real-time data feeds has become the low-cost alternative.

With the trend toward operational applications closely linked into the data warehouse, business requirements for low-latency data will continue to increase. Driven by "smart and fresh" data that is enabled by a new generation of enterprise applications, innovative business processes are emerging and gaining a strong presence in today's competitive organizations. Consequently, low-latency data aligned with new business processes is changing the complexion of business intelligence (BI) as enterprises cope with the rigors of the global economy. Acquiring low-latency data at a lower cost can propel these companies past their competition.

—R.H. and J.G.