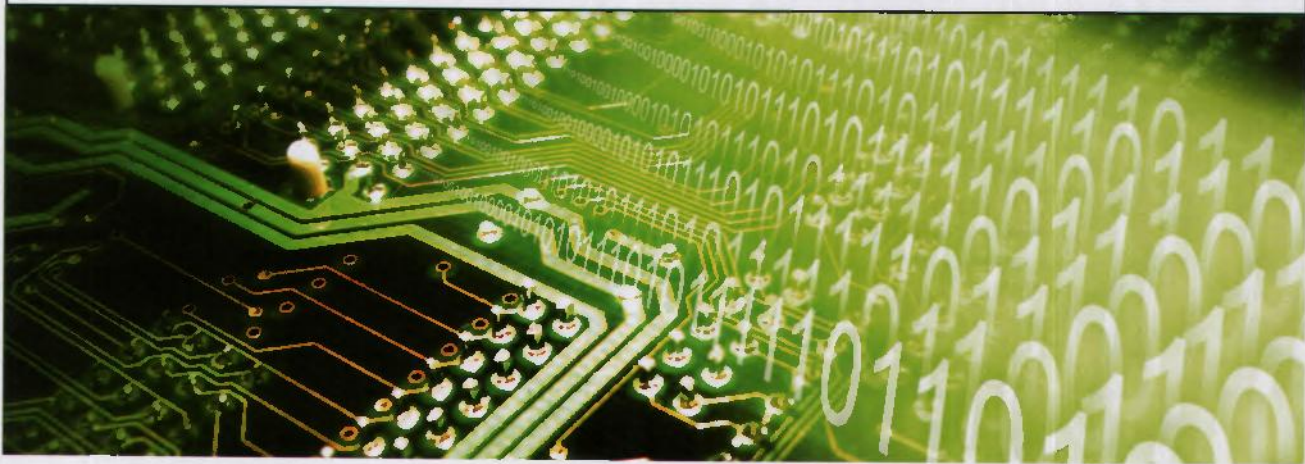




INSTANT GRATIFICATION

A real-time data warehouse may seem extravagant and unnecessary, but the evolution of business infrastructure may make it a necessity.



While most businesses aim to be responsive to their customers' needs, much of the planning that guides customer strategies has come from hours of careful analysis, where traditional business intelligence, based on the data warehouse, is supreme. But what if all that analytical power could be injected into customer interactions as they happened?

Today, users of an online banking system might expect to be targeted with offers based on historical transactions. For example, a user whose credit card is reaching its limit may trigger a loan advertisement popping up on screen. But a bank that is able to track the users actions in real time would be able to react instantly if it became obvious that during a session the customer's intention was actually to remortgage their house.

Such scenarios are not the fevered visions of some quack futurologist; they are happening today. The London Stock Exchange has deployed a state-of-the-art data warehouse that uses sub-second refresh rates to improve the depth and quality of the financial market intelligence services it delivers.

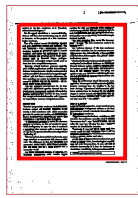
Mobile operator Orange is currently using a 24-hour update cycle for its Netezza data warehouse, but sees that refresh rate contracting ever tighter going forward.

"When we had a data warehouse that updated every few weeks, we could only use it for long term trend analysis," says Steve Hawkins, senior designer at Orange UK. "Now, we use it to provide more operational information, like the current status of a customer. Going forward, we will use the data warehouse increasingly as an operational tool". That move will put ever increasing demands on the data warehouse to update instantaneously.

However, in most cases business intelligence tools, and the data warehouses that provide the source material for that analysis, are used for high-level strategic reports and retrospective analysis. How can this be aligned with the desire to have real-time business intelligence? Indeed, is the effort involved even worth it?

"There is definitely a place for real-time access to information, but there is only a small subset of business applications that are truly real time in nature," says Andy Hayler, founder of data warehousing specialist Kalido. "Customers often perceive 'real-time' BI as something that is being pushed by the hardware vendors".

A recent survey by analyst group Gartner found that demand for lower latency in business intelligence is growing, however, the vast majority of businesses still find daily, rather than instantaneous,



updates to the data warehouse to be “practical, achievable and adequate.”

But the appeal of real-time, or more realistically, close to real-time data warehousing, may be about to increase, as the purpose of a data warehouse begins to change.

One of the original reasons for creating a separate data pool reserved for analysis and reporting – the data warehouse – was a technological limitation. Applying analysis to operational systems would ‘interrupt’ and impair the performance of operational processes that revolved around single applications such as financial ledger packages. Furthermore, meaningful analysis would need to draw data from many different systems.

However, today’s business processes typically span multiple applications, and this is increasing as business leaders choose to adopt a service-oriented architecture (SOA). Because these ‘composite applications’ pull data from a number of sources, the data repository can evolve to become the ultimate source of accurate data for the business – the so-called single version of the truth.

This has profound consequences for the data warehouse: operational systems demand that the warehouse is continuously updated; this updating process, and the analysis done on the data, must be fast enough to include information as it is generated, especially in customer facing circumstances.

DOUBLE TAKE

This dual purpose data warehouse, that serves both business analyst and frontline employee, is not merely a data warehouse with added processing power to let queries and updates run faster. Stephen Brobst, Teradata’s CTO, argues that the “active” data warehouse is a new hybrid form of database.

“A dimensional structure is needed for the strategic level of analysis, but this isn’t useful for operational intelligence,” he says. “So what is needed is a normalised data structure with a semantic layer on top that arranges the data dimensionally without duplicating it”.

This operational data-enabled warehouse in some ways resembles an operational data store, or data-mart. The difference is that the warehouse is a permanent, central record, rather than a point solution that stores data for a single purpose and is frequently wiped.

As Martin Richmond-Coggan, of corporate performance management tools vendor Applix

explains, the data mart approach to the problem of combining operational data and strategic analysis is only a temporary fix, and that having islands of data damages the very synchronicity that warehousing is designed to create.

“You can’t specialise [data marts] like that anymore”, he says. “The data warehouse needs to be business wide”.

The ‘real-time’ element of this data warehouse model refers to the fact that the warehouse can provide access to data, either directly or through some analytical filter, fast enough to form part of a front-line business process.

Whether that involves sub-second refresh rates, as used by the London Stock Exchange, or five minute intervals between updates should be determined by the business process, says Brobst. However, operational BI will certainly demand fast query and update speeds, and that means establishing far more stringent service levels than usually demanded of a data warehouse.

“SLAs for performance in real time are far more aggressive, because you’re talking about supporting business processes from the frontline to the top level”, he says. “Operational intelligence has a different customer base to standard BI. The analysis they do is not ad hoc and not exploratory, they are focused in the here and now. They’ve got a job to do and they want it to work quickly”.

DATA AS A SERVICE

As part of the SOA transition, many users have seen business process management software as the ideal mechanism to provide controls for the continuously updated data warehouse.

“What the companies that use a real-time or near real-time data warehouse are really doing is creating ‘data-as-a-service’”, says Teradata’s Brobst. But he adds this new way of using the data warehouse requires the functionality to evolve.

“Previously it was just the BI query tools that would access the data warehouse. Now it will be the point of integration of transactional services and decision services, via the BPM suite”.

As the focus of data warehouse evolves from past events to present states, its technological capabilities, and so for the mean time, its cost, will have to increase dramatically. It is time for businesses to decide how much living in the present is worth. ①

CONTACTS: pswabey@information-age.com