Find us on:

After the Fact is Not Good Enough

Posted By **Niek De Greef**
January 31, 2013

Tweet　0　　G+1　1　　　　　Share

A few months ago my debit card was skimmed (that is, copied illegally) at an airport in the US. While in transit from Amsterdam to North Carolina, I had withdrawn a sum of US dollars from my account at an ATM in the airport hall. A day and a half later my bank in the Netherlands called me and informed me they had blocked my debit card. They told me there were suspicious transactions on my debit card, and they didn't want to give further details or reveal whether the villains had managed to take money from my account.

How did the bank find out? It probably saw that I had used the card almost simultaneously in North Carolina and at some other place, where I couldn't possibly be at about the same time. But at what point would they have actually found suspicious behavior? What was the analytic and the data used to detect suspicious transactions? Could they have prevented possible malicious transactions?

The faster organizations can analyze the effect of business decisions or events, the bigger the advantage to the organization. If you can identify and prevent the fraudulent use of a credit or debit card at the moment a criminal executes a malicious transaction, this is clearly preferable to flagging a malicious event hours or days after the fact. Such an immediate detection allows you to deny this transaction and thus prevents the loss of money and image.

Numerous technologies are emerging to address the increasing need for a new category of near real-time analytics. Here I will examine one of these technologies that facilitates near real-time analytics in the domain of structured enterprise information.

Where existing analytical solutions fall short

There is a technical problem that limits IT solutions in supporting an analysis of near real-time enterprise data. What it boils down to is this: to shortcut many technical complexities, analytical computations need a different data layout than transactions to run efficiently. Furthermore, compared to transactional queries, analytical queries typically need a lot more computing resources.

Because of these different structural and operational characteristics, combining transactional workloads and analytical workloads has been very challenging if not practically impossible for many usage scenarios.

That is why today you find that the IT solutions built to support analytical and transactional workloads run on separate systems. Today's analytical solutions are typically built on dedicated servers. To feed these solutions with the necessary information, data has to be copied over from transactional systems and transformed into a data format optimized for analytics. As a consequence, the data used in the analytical solutions is not current, and the possibilities for using analytics in business transactions, such as for the identification of malicious transactions, are very limited.
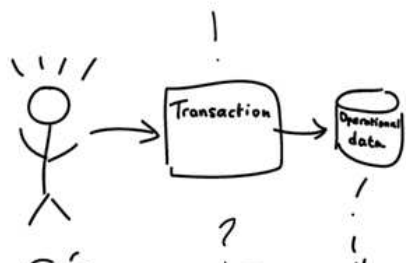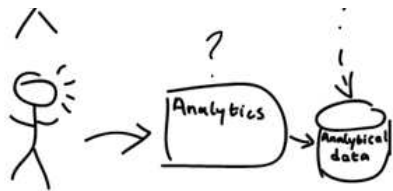
The marriage of transaction and analytics

Wouldn't it be nice if we could run analytical workloads on the actual, real-time information? A technology would be required that brings together the high-speed transactional capabilities and optimized analytical functions in a single solution. Such a solution would need to remove the data currency limitations of the analytical solutions and address the management burden of maintaining separate analytical systems.

The **IBM DB2 Analytics Accelerator** (IDAA) addresses these limitations. This appliance extends the DB2 database management software. A current copy of the data is held in the accelerator, in a format that is ideal for analytics, and the DB2 software takes care of managing that data in the accelerator all by itself.

Now analytical as well as transactional queries can be fired off to DB2. A piece of built-in intelligence in DB2 determines whether DB2 can best run the query natively or if it is more efficiently executed in the accelerator. Wherever executed, the resulting query is returned to the requestor through the DB2 interface. Thus the client that issues the query is not aware of the accelerator and can only suspect the existence of an accelerator by the speed through which analytical queries are run.

A fast lane braid: Acceleration

But it gets even better. The appliance uses special hardware facilities to run the analytical queries much faster than was possible before. The IDAA houses several technologies working together to speed up analytical queries. The accelerator implements a massively parallel computing model that is more often used for analytical solutions, but in this model the IDAA employs a number of unique technologies.

The IDAA facility opens up a whole new range of analytical possibilities. Not only is it now possible to run **operational analytics** on the near real-time information, having removed the need to copy the data to another system dedicated to analytics, but, just as important, it is now possible to include analytical operations in the transaction itself, making it possible to do *transactional analytics* that detects the malicious transaction at the moment it is attempted.
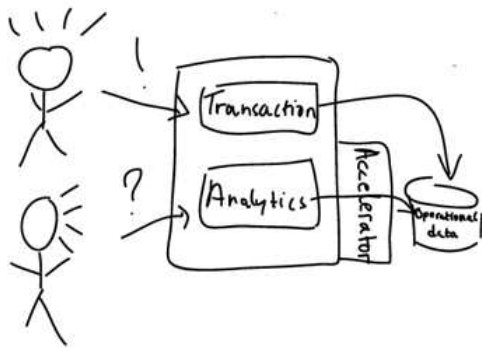
A data warehouse on steroids

It is great to be able to run analytical queries on near real-time data. And this only reduces the need for a traditional data warehouse. There are also use cases where having a data warehouse is a better solution. For example, there may be a need to combine and aggregate data from different IT solutions, such as **enterprise resource planning** (ERP) data with data from custom-built applications.

In the past, running a data warehouse on **DB2 for z/OS** was relatively expensive. The resource-intensive nature of analytical queries on data warehouses was a difficult marriage with the general-purpose architecture, aimed to run many different workloads in parallel, and the accompanying licensing scheme.

However, with the introduction of the IDAA the general purpose architecture of DB2 for z/OS is integrated with a special-purpose accelerator, and this combination provides a very cost-effective platform for data warehouses on DB2 for z/OS. As we have seen, the resource-hungry analytical data warehouse queries are offloaded to the accelerator. As a consequence this workload does not add to the general-purpose CPU cycles on which the DB2 database software charging is based.

So, the accelerator not only accelerates data warehousing on DB2 z/OS; it also improves the responsiveness.

More . . .

I could also mention the benefits of reporting on current data instead of data that is days old, or more. Or improved report availability. Or I could talk about improved data security and reduced risk of information leakage because data is managed in one place and not spread around.

Or I could discuss IT benefits like the avoidance of expensive extract, transform and load processes for filling business intelligence solutions. Or the improved system predictability and stability due to the reduced impact of long-running queries on transactional workloads. Or improved database administrator productivity because less tuning is required. Or about the storage savings you could make by archiving data with the high performance storage saver feature.

Maybe some other time.

The IBM DB2 Analytics Accelerator creates a hybrid platform for workload-optimized data workloads. As such it is an intriguing example of the hybrid nature of future enterprise systems.

Niek de Greef is an Executive IT Architect working for IBM in The Netherlands. Niek has more than 20 years of experience in IT. His areas of expertise include technology strategy, enterprise architecture, application integration, software engineering, and infrastructure architecture. You can reach him on Twitter **@NdeGreef1**.

Redbooks® thought leader

| Tweet | 0 | G+1 | 1 | | Share |

To effectively compete in today's changing world, it is essential that companies leverage innovative technology to differentiate from competitors. Learn how you can do that and more in the **Smarter Computing Analyst Paper from Hurwitz and Associates.**

Subscribe to the SC Blog »

This entry was posted in Big Data and Analytics and tagged analytics, DB2, Enterprise Systems, ERP, IDAA, real-time analytics, z/os. Bookmark the permalink.

Recent Posts

How do you build a system of insight?

Posted By Anirban Chatterjee
September 30, 2015

This week, IBM is making important strategic announcements at the Strata + Hadoop World conference in New York City. And in an important webcast on October 5, where Power Systems will unveil the next chapter in our journey to make superior systems of insight available to everyone.
Continue reading

Reach the forefront of digital transformation with waitless computing

Posted By Andy Dominguez
September 24, 2015

That's why IBM Power Systems continue to introduce products and services that make waitless computing a reality. On October 5th, tune in for the new advancements Power Systems is bringing to the market—and an entirely new way to consume Power Systems.
Continue reading

Comments

Logged in as fromthepolder   Dashboard | Edit profile | Logout

There are no comments posted yet. Be the first one!

Post a new comment

Posting as fromthepolder (Logout)

Subscribe to Replies   **Submit Comment**

Enter text right here!

We collect your name and email address solely for the purpose of accepting and posting your comment. Only your name will be posted with your comment, not your email address. Social media buttons on this site may log certain information such as your IP address, browser type and language, access time, and referring Web site addresses, and, if you are logged in to those social media sites, they may also link such collected information with your profile information on that site.