



Business Intelligence gets smarter - Q&A

Technology: Joanne Kelly 29 September 2005

Business intelligence (BI) is the silver bullet that renders overwhelming quantities of enterprise data manageable, meaningful and useful. Transforming data into knowledge is the aim of any BI application; at its best, BI empowers decision makers with deep real-time insight into the organisations finances, customers, inventory, supply chain and other critical business measures.

The forthcoming release of Microsoft SQL Server 2005 scheduled for Nov. 7 provides many new enhanced BI features designed to give users a competitive edge by increasing productivity while reducing costs. Though SQL Server 2005 is not scheduled for wide availability until November, Microsoft and a number of industry partners in the BI space have been working for months on a large-scale implementation based on actual customer scenarios to discover best practices for creating BI applications based on SQL Server 2005.

Project REAL (Reference implementation, End-to-end, At scale, and Lots of users) is an innovative exercise in system development, deployment and testing. It provides lessons in analyzing large data sets that are applicable to many customers seeking powerful BI tools. The project is already well documented on the Web at <http://www.microsoft.com/sql/bi/ProjectREAL> and will be demonstrated in a series of hands-on sessions at this weeks 2005 PASS (Professional Association for SQL Server) Community Summit, the largest event of the year exclusively dedicated to SQL Server education.

Below are PressPass questions with Bill Baker, general manager of SQL Server Business Intelligence at Microsoft and the keynote speaker at this years PASS Community Summit, about Project REAL and putting business intelligence into practice.
What is the purpose of this new initiative?

Baker: Project REAL demonstrates that business intelligence in SQL Server 2005 is ready to use. Customers can take the tools in SQL Server 2005, use them as we have and achieve success in large-scale operations. Were providing a road map for how to build systems like this. We built the system so that customers can learn from it. Were going to be talking a lot about it, publishing papers and show casing the system at a number of conferences beginning with the PASS Conference. The whole idea is that weve developed best practices and a great deal of supporting information so that customers can be successful with BI applications at the kinds of scales and complexities with which were working in Project REAL.

What prompted you to take this approach?

Baker: We knew that there were many important changes being made in the BI tools in the SQL Server 2005 release. We wanted to get out ahead of the curve and update best



practices. The best way for us to understand where we needed to develop new guidelines was to go through the process of building and operating a large-scale system ourselves.

The acronym Project REAL for Reference implementation, End-to-end, At scale and Lots of users really captures the idea that this is a real-world implementation. We took Barnes & Noble's data warehouse and literally rebuilt it on SQL Server 2005, using all the latest tools and techniques. And we now have a real-world, real-complexity, large-scale implementation that we can show our customers. Being able to show what we did and how we did it is a key aspect of this project. That's the reference implementation piece.

How is the Barnes & Noble data store representative of the data sets of other business customers?

Baker: That's a really important question because there are a number of ways in which Barnes & Noble mirrors many customer scenarios. One is that they bring data from multiple sources and need to integrate the data into one data warehouse. Barnes & Noble brings some data from a sales tracking system and other data from an inventory system. And they're two different systems. They pull this data in, integrate it and build one relational database, which is the interior of their data warehouse.

The second thing about Barnes & Noble is the disparate audiences they have for their data. They want to support managers out on the front lines, working in the stores. They also want to support buyers responsible for managing the supply of products over their whole chain. And they want to support senior managers in the New York headquarters who are asking very different kinds of questions than those being asked by store managers out in the field. All of these different user populations are being supported from the same data warehouse. We find that's a very common scenario in the real world these days: many communities of users must be supported from one data warehouse.

A third, really important way in which Barnes & Noble is representative relates to challenges of combining sales and inventory data. If you look at traditional OLAP systems and Analysis Services is an example of an OLAP system they tend to have a really difficult time with inventory data. That's a class of data that we call semi-additive measures.

What's the problem with that particular type of data?

Baker: I'll give you an example. Contrast inventory data with sales data. If I sell five copies of a book in January and five more in February and five more in March, then I can safely say that in the first quarter, I sold 15 copies of that book. Because it adds up naturally.

Inventories don't work that way. If I had five copies in inventory at the end of January and five copies in inventory at the end of February and five in inventory at the end of March, I wouldn't say I had fifteen copies in inventory. It doesn't add up. There are certain classes of real world issues that don't add up like that. Inventory is a big, canonical example of semi-additive measures.

So that was one of the big challenges you faced in taking all the Barnes & Noble data and



applying BI in the new version of SQL Server 2005?

Baker: It really wasn't a challenge at all. That's what's so interesting. Analysis Services now deals with semi-additive measures very naturally. It's simple and easy to do. We're showing that working very, very well in Project REAL.

You're working closely with a number of partners on this initiative. Why was it conceived as a cooperative effort?

Baker: When we got into this a little over a year ago, one of our goals was to broadly disseminate the knowledge that would come out of the work. And we felt that a way to do that was to engage with several best-of-breed partners who would be able to use this information and communicate it in their own work and practices. We sought out partners in the BI space who were really willing to contribute. Everybody did real work on this project. We've had a great time working these partners: Apollo Data Technologies, EMC, Intellinet, Panorama, Proclarity, Scalability Experts and Unisys. And, of course, Barnes & Noble was very generous in providing the source data set.

What kinds of business scenarios are you running in the project?

Baker: This is full-scale effort to apply our BI tools to the entire data warehouse. That's an important notion the 'e' in Project REAL is for "end-to-end." We're covering front-to-back, everything that a customer has to step up to build in a large-scale data warehouse. They need to do ETL from their source system into their relational warehouse. They need to build cubes on top of the warehouse. They want to build reports on top of the cube and the relational database. They want to deploy applications out to their users' desktops, so they need to be able to deploy Web applications and work with partners across the Internet. End-to-end is a really important aspect of this. And so we touch all the components of the BI stack: Integration Services, Analysis Services, Reporting Services and the relational engine as well.

How will Project REAL benefit customers?

Baker: We expect that there will be two primary take-aways. First, what we're doing here will help customers with their decision-making process. When they're considering deployment of a data warehouse, having a roadmap they can follow will make it easier to decide to go ahead with the project and will provide tremendously valuable information on how to architect the project.

Second are the best practices and guidelines we're providing for developers and operations people. These are the folks doing the day-to-day work of deploying and running these systems. They can read about how we made it all work, what we learned and many recommendations we make. What to do and what not to do. We're publishing a phenomenal volume of information in white papers on our Web site. We've got nine papers posted now with more to come. In addition, we're doing webcasts and providing a number of live educational opportunities this week at the PASS Conference and in November at the SQL Connections Conference in Las Vegas.

What is the state of the art in BI?



Baker: Business intelligence is becoming ubiquitous and its happening in a number of different ways. One is that we're seeing the deployment of reports out of BI systems into all different parts of the company, at every level. This is no longer being used only by the business analysts and senior managers. These tools are now having an impact all throughout the organization.

The state of the art involves enabling use of the data by all parts of the organization, having it be current and relevant to people in their various roles. To present it in ways that are naturally useful to people: What's useful to the business analyst is not useful to the operator on the telephone. And to be able to explore data in innovative ways, not only OLAP-type analysis but through data mining as well. And to see all of this embedded in the day-to-day operations of the company.

How is Microsoft helping customers get there?

Baker: If you look at every one of the components that we're shipping, they address different aspects of that ubiquity promise. Integration Services addresses the question of synthesizing data from multiple sources into relevant and timely information. In Analysis Services, both the OLAP and data mining pieces are all about doing powerful analytics on data so you can understand it in new ways that you may never have thought of before. Reporting Services is about getting that information out to all parts of the company and making it useful by a wide variety of people in forms that they are able to use naturally in their roles. And, of course, the relational engine is what stands behind it all and is the underpinning of the entire system.

What have you learned through the performance-testing approach you've taken with Project REAL?

Baker: We're shipping a great platform. In many ways I'm just really, truly blown away by the richness of the platform. As my team and our partners worked together building this representative system, every day we saw new ways to do things and new opportunities. I'm really impressed with the product. The Project REAL team will be at the PASS conference all week in the Microsoft booth. We're looking forward to discussing this project at Ask the Experts sessions or simply when someone cares to approach one of us between sessions.

With all the excitement building around the November release of SQL Server 2005, what is your thinking about the importance of this release and its value to customers?

Baker: This release is going to change the world of BI. Every piece of the product has got new functionality. Integration Services takes enterprise-grade ETL and presents it in a package that's accessible to everybody. Analysis Services brings new classes of analysis that weren't possible before, such as attribute analysis, semi-additive measures and many-to-many relationships. There are now whole new classes of questions that it's possible to ask and get answered that weren't possible before. And of course with all the new data mining algorithms and the integration of data mining with Analysis Services, Integration Services and Reporting Services, we will be making that incredibly powerful tool available to audiences that would have never dreamed of having it before. And while all of what we're providing in BI with SQL Server 2005 is of mission-critical importance to



enterprise-scale businesses, we see this as having real value to every organization. There's something for everyone in SQL Server 2005.