

DATABASE

TRENDS AND APPLICATIONS

Solutions for the Information Project Team • www.dbta.com

Volume 18, Number 7 • July 2004

Good Data Modelers Make Safe (Ware)houses

By **Bill Laurent**

When embarking on a data warehouse project, the team assembled will be responsible for diverse missions, from learning the core business to ferreting out the details of the current IT system's topology -- the architectural point of departure. The complexity and size of even the smallest data warehouse projects will stretch existing resources. A formidable number of tasks will be assigned to team members based on their strengths, personalities, availability, and other intangibles. As the data warehouse battle is joined, you may share the trenches with various allies such as data architects, programmers, business analysts, project managers, network engineers, database administrators (DBAs), and my closest of confederates--data modelers.

The best data modelers are hands-on technologists (preferably DBAs) blessed with a business understanding of the subject matter to be modeled. Having an understanding of the business facilitates a higher quality of design, as well as faster information and requirements gathering, with primary business objectives and values held in esteem. Being cut from the DBA mold will reap enormous benefits downstream as the data modeling process will iteratively take into account things like table growth, locking issues, disk striping, index strategies, object placement, and other hands-on architectural and logistic requisites. The days of the pure data modeler are waning rather quickly: Companies seek additional value from modeling experts as the business comes to the realization that no modeling should be done in a vacuum. A trained business and technical eye directed

towards physical model implementation is imperative.

Data models and systems/network architectural blueprints are among the first components to be tackled in re-writing or building from scratch a data warehouse, business intelligence (BI) system, or reporting repository. If these critical constructs are flawed, the project is likely doomed to failure from the beginning. A logical model that fails to properly capture, explain, and promulgate business rules is bad; however, a poor physical fulfillment of the best logical models could spell greater long and short-term disaster.

Think of the value and advantages added to a project when the people who are tasked with designing logical data models are also capable of realizing the model's physical form as well. I can hear the arguments now: "The physical model is always going to be different from the logical model"; "The two are separate, and the physical model may vary greatly"; "Logical design must conform exactly to the business, and these models should never be created with an eye toward their physical implementation." I have never seen a smoothly run project where a logical modeler turned over his model to a DBA without heavy--sometimes impassioned, always costly in time--discussions and back-and-forth Q&A sessions. While having a business content expert create a logical model that mirrors exactly the business practice is usually a "nice to have;" it will seldom have a large captive audience. Time and resource constraints often dictate that the logical model is the physical model in most respects.

It is important for the data warehouse modeler to know the physical implica-

tions of his/her original logical model and how they fit in with the entire warehouse infrastructure. Models will have to constantly bend to accommodate the usual moving targets--from development objectives to project timelines to reporting and data-set distribution requirements--as well as physical hardware limitations; risk must constantly be weighed, evaluated, and managed on a daily basis. Ongoing difficult judgments and compromises in normalization, disk partitioning, object-to-segment/disk placement, data redundancy and storage, and other server side particulars will have to be made with wisdom and experience. Not only will the best data modelers be DBA-conversant, they will have many of the unmistakable talents of a data architect and will always be thinking about architecture, network, and deployment issues; knowing the power, processors, and capability of your servers, and where every data mart and database object fits into the big and small picture.

Good data modelers will continually be able to make educated guesses about data distribution, storage, batch limitations, performance, and the administrative overhead involved to horizontally and vertically manage the physical implementations of logical tables, during all phases of the modeling process. It will be important to staff your projects with the most effective resources, hence you consider wisely the total skill-set of your data modelers.

An expanded version of this article appears at www.dbta.com.

Bill Laurent is the executive vice president of Loyer TCG. He can be reached via e-mail at blaurent@loyertcg.com.