www.kimballgroup.com

Number 15, October 29, 2000

Kimball Design Tip #15: Combining SCD Techniques

KIMBALL GROUP

By Margy Ross

The acronym, SCD, is a keyword in a dimensional modeler's vernacular. As most of you know, SCD is short-hand for slowly changing dimensions.

KIMBALL University

There are several well-documented techniques for dealing with slowly changing dimension attributes. Briefly, with SCD Type 1, the attribute value is overwritten with the new value, obliterating the historical attribute values. For example, when the product roll-up changes for a given product, the roll-up attribute is merely updated with the current value. Using Type 2, a new record with the new attributes is added to the dimension table. Historical fact table rows continue to reference the old dimension key with the old roll-up attribute; going forward, the fact table rows will reference the new surrogate key with the new roll-up thereby perfectly partitioning history. Finally, with Type 3, attributes are added to the dimension table to support two simultaneous roll-ups - perhaps the current product roll-up as well as "current version minus one", or current version and original.

In my experience, data warehouse teams are often asked to preserve historical attributes, while also supporting the ability to report historical performance data according to the current attribute values. None of the standard SCD techniques enable this requirement independently. However, by combining techniques, you can elegantly provide this capability in your dimensional models.

We'll begin by using the SCD workhorse, Type 2, to capture attribute changes. When the product roll-up changes, we'll add another row to the dimension table with a new surrogate key. We'll then embellish the dimension table with additional attributes to reflect the current roll-up. In the most current dimension record for a given product, the current roll-up attribute will be identical to the historically accurate "as was" roll-up attribute. For all prior dimension rows for a given product, the current roll-up attribute will be overwritten to reflect the current state of the world. If we want to see historical facts based on the current roll-up structure, we'll filter or summarize on the current attributes. If we constrain or summarize on the "as was" attributes, we'll see facts as they rolled up at that point in time.

We've described a hybrid approach that combines the three fundamental SCD techniques. We're creating new rows to capture change (Type 2), adding attributes to reflect an alternative view of the world (Type 3), which are overwritten for all earlier dimension rows for a given product (Type 1). As a student recently suggested, perhaps we should refer to this as Type 6 (2+3+1)...