**Consejo de diseño 137 Crear y administrar dimensiones reducidas.**

Este consejo de diseño continua mi serie de implementar patrones de diseño comunes ETL. Estas técnicas deben probar valor para todos los desarrolladores de sistemas ETL, y, esperamos, proporcionar algunas guías de producto también para las compañías de software ETL.

Renombrar esto como una dimensión reducida es un subconjunto de un atributo dimensional que aplica en un nivel superior de recopilación. Por ejemplo, una dimensión mensual seria una dimensión reducida de la dimensión fecha. La dimensión de mes se podría conectar con una tabla de hechos pronosticados cuyo granulo se sitúa en el nivel mensual.

Recientemente, me he topado con un buen ejemplo de la necesidad de una dimensión reducida de un entusiasta del Kimball group que trabaja para una compañía que gestiona propiedades de centro comercial. Recogen algunos hechos en el nivel de almacenaje como pagos de los alquileres, y otros hechos en general al nivel de la propiedad como afluencia de clientes y los costes de utilidad. Recordad, un objetivo fundamental de diseño es recoger datos desde el menor nivel de granulo. En este caso, primero intentaremos asignar los datos del nivel de propiedad hasta el nivel de tienda. Sin embargo, la compañía en cuestión pensó que algunos datos sobre la propiedad no podrían estar asignados al nivel de tienda; para ello necesitarían tablas de hechos en ambos niveles la tienda y la propiedad. Esto significa que también necesitaban dimensiones en el nivel de la propiedad.

**Crear la dimensión base**

Hay muchos modo de crear una dimensión reducida, dependiendo de cómo se estructuran los datos en el sistema fuente. El camino más fácil es crear la dimensión base primero. En este caso, construir la dimensión Almacén extrayendo las claves naturales de los niveles de tienda y propiedad y los atributos de la fuente, asignando claves subrogadas y rastreando los cambios de los atributos importantes con seguimiento de cambio tipo 2.

La dimensión Almacén tundra diversos niveles de atributos de propiedad incluyendo la clave natural de la propiedad porque los usuarios querrán deslizarse por los hechos de tienda utilizando las descripciones de propiedad. También realizarán preguntas que sólo impliquen la relación entre tienda y propiedad, como “¿Cuál es la media de tiendas por propiedad?

**Crear la dimensión reducida desde la dimensión base**

Una vez que la dimensión del nivel más bajo está en su sitio, crear la dimensión reducida inicial, en este caso la dimensión de propiedad, es esencialmente lo mismo que crear la mini dimensión que describimos en el consejo de diseño 127. Identifica los atributos que necesitas extraer de la dimensión base y crea una nueva tabla con una columna de clave subrogada. Rellena la tabla utilizando SELECT DISTINCT de las columnas de la dimensión base junto con un campo IDENTITY o SEQUENCE para crear la clave subrogada. En el ejemplo de la propiedad, el siguiente SQL te serviría para empezar:

INSERT INTO Dim\_Property  
SELECT DISTINCT Property\_Name, Property\_Type, Property\_SqFt, MIN(Effective\_Date), MAX(End\_Date)  
FROM Dim\_Store  
GROUP BY Property\_Name, Property\_Type, Property\_SqFt;

El procesamiento incremental es un poco más desafiante. El enfoque más fácil si trabajas desde una dimensión base existente como hemos descrito es usar el método de la fuerza bruta. Crea una dimensión reducida temporal aplicando el mismo SELECT DISTINCT a nueva dimensión base cargada. Entonces procesa cualquier cambio de tipo 2 comparando las filas actuales con la dimensión reducida temporal (WHERE End\_Date = ‘9999-12-31’) con las filas actuales de la dimensión reducida maestra basada en la clave natural de la dimensión reducida. Puede sonar ineficiente, pero realizar un test sólo lleva 10 segundos en contraposición a una dimensión base de alrededor de 1.1 millones de filas dentro de una máquina virtual de un ordenador portátil.

**Alternativa: Crea la base y la dimensión reducida separadamente.**

Si tienes tablas fuente separadas para los atributos de las dimensiones reducidas, puedes utilizarlas para crear las dimensiones reducidas directamente. De nuevo, asigna las claves subrogadas, y rastrea los cambios utilizando el seguidor de cambios de tipo 2.

La dimensión reducida completa puede ser unida de nuevo con la dimensión base en el proceso ETL para rellenar los atributos actuales de nivel superior. Entonces puedes proceder con el seguimiento de cambios en la dimensión base. La parte difícil es crear la base histórica y las dimensiones reducidas porque tienes que compara ambas las fechas efectivas y las fechas finales de ambas dimensiones, insertando nuevas filas en la dimensión base cuando se den los cambios en la dimensión reducida.

**Presenta las dimensiones a los usuarios**

Una vez que tienes tus dimensiones históricas llenas y los procesos ETL en su lugar, la decisión final es como representar estas tablas al usuario. La dimensión reducida puede ser presentada exactamente como existe, unida a hechos de nivel superior tal y como se necesite. Cuando los usuarios ven la dimensión base, les gusta ver los atributos asociados de las dimensiones reducidas también. Esto puede hacerse de diversas maneras; ya sea considerando la unión de la dimensión base con la dimensión reducida a través de la clave subrogada de la dimensión reducida, o definiendo las uniones en la herramienta BI de tu capa semántica, o realmente uniendo las tablas en el proceso ETL y físicamente instanciando a las columnas de dimensión reducida en la dimensión base.

Todos estos enfoques deben parecer exactamente el mismo para los usuarios. La única diferencia estará en la ejecución, y tendrás que experimentar con tu sistema para averiguar qué es lo que agiliza la mayoría de las consultas. Normalmente nos damos cuenta de que encontramos la mejor ejecución y la capa semántica más simple pre-uniendo las tablas del proceso ETL y copiar físicamente las columnas, a menos que sean tablas muy largas,

<http://www.kimballgroup.com/2011/08/design-tip-137-creating-and-managing-shrunken-dimensions/>