**Consejo de diseño 147 Claves “Súper naturales” duraderas.**

Una de las tareas de la gestión del sistema de dimensión del cliente es “asignar una única clave duradera a cada cliente”. Por clave duradera, entendemos un único valor clave que de manera única y confiable identifique un cliente dado en el tiempo. En la mayoría de los casos, esta clave única duradera es la clave natural del negocio obtenida desde el sistema operacional, y todo lo que tenemos que hacer es copiarla como un atributo en la tabla dimensional. Sin embargo, hay casos donde las claves naturales cambian, y cuando esto sucede, el administrador del sistema tiene que intervenir.

Algunas causas comunes que justifican los cambios en las claves naturales incluyen razones empresariales, entradas duplicadas e integración de datos desde múltiples fuentes. Todo esto requiere la creación y administración de una única clave duradera, también conocida como clave súper natural, durante el proceso ETL.

Un buen ejemplo de cambio en una clave natural debido a razones empresariales proviene de la industria de las tarjetas de crédito. Los números de cuenta de las tarjetas de crédito son claves naturales; aparecen en las transacciones y se mapean con las claves subrogadas en la tabla dimensional. Si una tarjeta de crédito es robada, un nuevo número de cuenta se emite. Sin darse cuenta de este cambio, el nuevo número de cuenta aparecería como un número totalmente nuevo y se introduciría como una nueva entidad en la dimensión de cuenta. El historial completo de la cuenta se perdería porque ahora tendría dos claves naturales.

Cuando este tipo cambios dirigidos por el negocio suceden, el sistema de transacciones debe generar un informe de notificación diciéndole al proceso ETL que una nueva cuenta ha sido creada para reemplazar una cuenta antigua. Esto podría ser tan simple como una tabla con el número antiguo de cuenta, el número nuevo, y la fecha efectiva. El sistema ETL debe entonces crear una nueva fila con el nuevo número de cuenta y una columna de la clave duradera separada que ligue las cuentas antiguas y las nuevas juntas. La figura 1 muestra como esta clave duradera aparecería en una dimensión de cuentas.

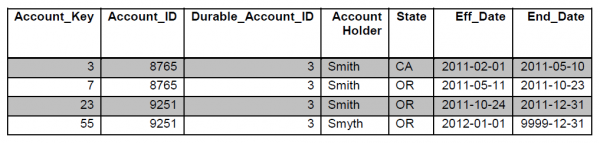
[](http://www.kimballgroup.com/wp-content/uploads/2012/07/design-tip-147-fig-1-e13436640286961.png)

Figura 1 – Una clave duradera en una dimensión de cuenta.

En la figura 1, la clave de cuenta es la clave subrogada asignada al sistema ETL para identificar únicamente a cada fila. La cuenta \_ID es el sustituto ETL para la clave natural del sistema de transacciones porque normalmente no cargarías a un elemento sensible como una tarjeta de crédito un número de cuenta directamente desde el data warehouse. La tercera columna, llamada Durable\_Account\_ID, es la clave duradera asignada por el sistema ETL para enlazar todas las filas relacionadas juntas. La figura q muestra cuatro filas para la misma cuenta porque había dos cambios de tipo 2 para seguir la trayectoria, un cambio de estado y un cambio de apellido, además del cambio de Account\_ID.

Otro patron de diseño útil es añadir la clave de cuenta duradera a la tabla de hechos además de la clave subrogada de la dimensión. Esto se une de nuevo a las filas actuales en la dimensión para facilitar el informe de toda la historia mediante los atributos de dimensión actuales. Consultad el artículo titulado “ Slowly changing dimensions no siempre son tan fáciles como 1,2,3” para más información.

Ocuparse de las entradas duplicadas en una dimensión, o integrar las fuentes dispares en una dimensión única, implica una lógica de negocios compleja que depende de claves duraderas. El resultado final es similar a la Figura 1, pero el proceso de integración tiene que generar la lista de elementos relacionados más que relacionarse con el sistema de transacciones. La figura 2 muestra las integraciones de los productos desde múltiples sistemas fuente. El subsistema MDM en el proceso ETL identifica estos tres productos como el mismo y les asigna la misma clave duradera.

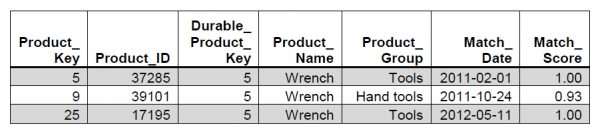
[](http://www.kimballgroup.com/wp-content/uploads/2012/07/design-tip-147-fig-2-e13436642838501.png)

Figura 2 – Entradas duplicadas de producto

En este caso, necesitas diferenciar las claves naturales de las distintas fuentes para que no choquen entre ellas. Intenta utilizar un tipo de característica de datos para la clave natural de origen, añadida por un código fuente. Por ejemplo, si los productos de la Figura 2 vienen del sistema SAP o el sistema CRM, la columna de Product\_ID contendrá los siguientes valores: SAP|37285, CRM|39101, y SAP|17195. Esta tabla se convierte en input de la dimensión y administradores de hechos.

Las claves duraderas son obligatorias para tratar con ambigüedades en las claves naturales del sistema fuente. Crear y asignar claves duraderas te permite trabajar en torno a los cambios del negocio con las claves naturales, o integrar, duplicar o separar datos. Pero las claves duraderas son sólo el principio; hay mucho más sobre deduplicación y integración de datos.

<http://www.kimballgroup.com/2012/07/design-tip-147-durable-super-natural-keys/>