**Consejo de diseño#145 Tablas de hechos de períodos de acumulación de instantáneas.**

En el consejo de diseño 140, discutía los desafíos de diseñar esquemas dimensionales para procesos de longitud indeterminada como el flujo de ventas o el procesamiento de quejas de seguros. Concluimos que se presentan mejor como tablas de hechos de instantáneas acumuladas caracterizadas por una fila por flujo existente cuando cada línea se actualiza múltiples veces a lo largo de su ciclo de vida. Sin embargo, debido a que cada fila se actualiza, tenemos un registro imperfecto de la historia. Las instantáneas acumuladas hacen un gran trabajo al decirnos el estado actual de los flujos, pero pasan por alto los estados intermedios. Por ejemplo, una queja que pueda moverse dentro y fuera de los estados múltiples de veces: abierto, denegado, criticado, re-abierto, re-cerrado. Las instantáneas acumuladas son enormemente valiosas, pero hay diversas cosas que no pueden hacer:

* No nos pueden mostrar detalles de cuándo y porque la queja pasó por algún estado múltiples veces.
* No podemos recrear nuestro “libro de negocios” en una fecha arbitraria en el pasado.

Para resolver estos dos problemas, necesitaremos dos tablas de hechos. Una tabla de hechos transaccional recoge los detalles de los cambios individuales de estado. Entonces añadiremos fechas efectivas y de caducidad a las instantáneas acumuladas para registrar su historia.

La tabla de hechos transaccional es sencilla. Normalmente emparejamos la tabla de hechos de instantáneas acumuladas con una tabla de hechos transaccional que contiene una fila para cada cambio de estado. Donde la instantánea acumulada tiene una fila por proceso como una queja, la tabla de datos transaccional tiene una fila por suceso. Dependiendo de tu sistema fuente, es común construir la tabla transaccional primero, y derivar la instantánea acumulada de ello.

Ahora vamos a centrar nuestra atención en la tabla de hechos de períodos de acumulación de la instantánea. Primeramente, no todo el mundo necesita preocuparse por retener estas instantáneas marcadas en el tiempo. Para la mayoría de organizaciones, una instantánea estándar acumulada representando el estado actual del flujo, combinada con la tabla de hechos transaccional para mostrar los detalles del suceso, es suficiente. Sin embargo, hemos trabajado con diversas organizaciones que necesitan entender la evolución de un flujo. Mientras que es técnicamente posible hacerlo desde la transacción de datos, no es un juego de niños.

Una solución al requisito de rastreo del flujo histórico es combinar las instantáneas acumuladas con instantáneas periódicas: saca una fotografía del flujo en un intervalo regular. Este método de fuerza bruta es excesivo para los flujos que tienen una duración relativamente larga, pero cambian frecuentemente. Lo que mejor trabaja en este caso es añadir el seguimiento de cambios en efectivo y caducidad a las instantáneas acumuladas.

Aquí está como funciona:

* Diseña una tabla de hechos estándar de las instantáneas acumuladas.
* En lugar de actualizar cada fila tal y como cambia si estado, añade una nueva fila. Nuestros diseños más recientes se han realizado hasta el granulo diario: añade una nueva fila a la tabla de hechos cualquier día cuando algo sobre el flujo haya cambiado (p.e, quejas, proceso de ventas, o reacción contra las drogas).
* Necesitas columnas adicionales de metadata, similares a una dimensión de tipo 2:
  + Instantánea del día de comienzo: la fecha en que esta fila se hizo efectiva.
  + Instantánea del día de finalización: la fecha en que esta fila expira, actualizada cuando una nueva fila se añade.
  + Instantánea del indicador actual: actualizada cuando añadimos una nueva fila cuando ocurre este flujo.

La mayoría de usuarios solo están interesados en la visión actual, p.e, instantánea acumulada estándar. Puedes conocer sus necesidades definiendo un criterio ( probablemente una perspectiva indexada o materializada) que filtre las filas de instantáneas históricas basadas en el indicador de instantánea actual. Alternativamente, debes elegir iniciar una tabla física de filas actuales al final del ETL de cada día. La minoría de los usuarios e informes que necesitan revisar el flujo de cualquier fecha arbitraria en el pasado puedo hacerlo fácilmente filtrando la instantánea de fechas inicial y final.

La tabla de hechos de períodos de acumulación de instantáneas es ligeramente más complicada de mantener que una instantánea estándar, pero la lógica es similar. Donde la instantánea acumulada actualice una fila, la instantánea marcada en el tiempo actualiza la fila conocida como actual e inserta una nueva fila. La gran diferencia entre la estándar y la instantánea marcada en el tiempo es el recuento de filas en la tabla de hechos. Si una queja media se cambia 20 días durante su vida, la instantánea marcada en el tiempo será veinte veces mayor que la instantánea estándar acumulada. Echa un vistazo a los datos de los requerimientos de tu negocio para ver si tiene sentido para ti. En nuestros diseños recientes, nos hemos sorprendido gratamente por lo eficiente que es este diseño. Aunque se cambiaron unos pocos flujos problemáticos cientos de veces, la gran mayoría se mantuvieron y se cerraron con un modesto número de cambios.

<http://www.kimballgroup.com/2012/05/design-tip-145-time-stamping-accumulating-snapshot-fact-tables/>