

Inmon vs Kimball

First Author

Esperilla Ruiz Adnner

adnnerer@hotmail.com

Abstract

1 Introduction

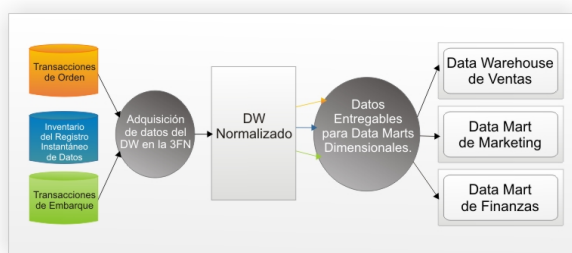
2 Objetivos

1. Determinar la diferencia entre las dos perspectivas presentadas.

3 Marco Teórico

3.1 El enfoque de Inmon

1. **Orientado a temas:** Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí.
2. **Integrado:** La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes.
3. **No volátil:** La información no se modifica ni se elimina, una vez almacenado un dato, éste se convierte en información de sólo lectura, y se mantiene para futuras consultas.
4. **Variante en el tiempo** Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones.



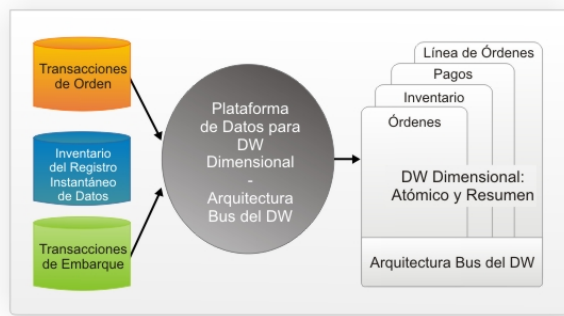
El enfoque Inmon también se referencia normalmente como Top-down. Los datos son extraídos de los sistemas operacionales por los procesos ETL y cargados en las áreas de stage, donde son validados y consolidados en el DW corporativo, donde además existen los llamados metadatos que documentan de una forma clara y precisa el contenido del DW. Una vez realizado este proceso, los procesos de refresco de los Data Mart departamentales obtienen la información de él, y con las consiguientes transformaciones, organizan los datos en las estructuras particulares requeridas por cada uno de ellos, refrescando su contenido.

La metodología para la construcción de un sistema de este tipo es la habitual para construir un sistema de información, utilizando las herramientas habituales (esquema Entidad Relación, DIS (Data Item Sets, etc). Para el tratamiento de los cambios en los datos, usa la Continue and Discrete Dimension Management (inserta fechas en los datos para determinar su validez para las Continue Dimension o bien mediante el concepto de snapshot o foto para las Discrete Dimension).

3.2 El enfoque Ralph Kimball.

El Data Warehouse es un conglomerado de todos los Data Marts dentro de una empresa, siendo una copia de los datos transaccionales estructurados de una forma especial para el análisis, de acuerdo al Modelo Dimensional (no normalizado), que incluye, como ya vimos, las dimensiones de análisis y sus atributos, su organización jerárquica, así como los diferentes hechos de negocio que se quieren analizar.

Los diferentes Data Marts están conectados entre sí por la llamada bus structure, que contiene los elementos anteriormente citados a través de las dimensiones conformadas (que permiten que los usuarios puedan realizar queries conjuntos sobre



los diferentes data marts, pues este bus contiene los elementos en común que los comunican).

La metodología para la construcción del Dw incluye las 4 fases que vimos en la entrada anterior del blog, que son: Selección del proceso de negocio, definición de la granularidad de la información, elección de las dimensiones de análisis e identificación de los hechos o métricas. Igualmente define el tratamiento de los cambios en los datos a través de las Dimensiones Lentamente Cambiantes (SCD).

4 Análisis

¿Cómo afectan estos factores a los dos modelos de datawarehouse?

Factores

1. Presupuesto
2. Plazos
3. Expertise
4. Alcance
5. Mantenimiento

Inmon

1. Coste inicial alto
2. Requiere más tiempo de desarrollo
3. Equipo con especialización alta
4. Toda la compañía
5. Fácil mantenimiento

Kimball

1. Coste inicial bajo
2. Tiempo de desarrollo inferior

3. Equipo con especialización media

4. Departamentos individuales

5. Mantenimiento más complejo

Aparte de estas variables, que van a ser determinantes a la hora de decidimos por un modelo u otro, tenemos que tener muy clara cuál será la finalidad de nuestro datawarehouse.

Conclusiones

El datawarehouse de Kimball está orientado a la consulta de la información, por lo que su estructura interna está especialmente diseñada para garantizar una explotación de los datos rápida y sencilla, no requiriendo usuarios especializados para ello. Por el contrario, el datawarehouse de Inmon persigue la integración de todos los datos de la compañía, estando orientado hacia el almacenaje de grandes volúmenes de datos, por lo que su estructura interna normalizada se diseña para evitar la redundancia de datos, simplificar las labores de mantenimiento, etc. cuestiones que complican las consultas de la información, requiriendo que los usuarios finales estén mucho más especializados. Así, podríamos decir que el enfoque de Kimball se ajusta más a proyectos pequeños en los que se persiga un sistema fácilmente explotable y entendible por el usuario y de rápido desarrollo, siendo el modelo de Inmon más apropiado para sistemas complejos de mayor envergadura. Todo proyecto tiene sus propias peculiaridades, siendo cada caso único e independiente, por lo que resulta necesario llevar a cabo un estudio de todas ellas antes de decantarnos por una solución u otra, de forma que podamos hacernos una idea sobre qué modelo se ajusta mejor a las condiciones de nuestro proyecto.

References

churriwifi <https://churriwifi.wordpress.com/2010/04/19/15-2-ampliacion-conceptos-del-modelado-dimensional/>.

bi-geek

<https://blog.bi-geek.com/arquitectura-comparativa-inmon-y-kimball/>