

INMON VS KIMBALL

Panty Sihuayro, Juan Carlos, Huilca Umpiri, William Arturo, and Esperilla Ruiz, Adnner Sleyder
Universidad Privada de Tacna \Facultad de Ingenieria \Escuela Profesional de Ingenieria de Sistemas

Resumen

Los almacenes de datos (data warehouses en inglés) toman cada día mayor importancia, a medida que las organizaciones pasan de esquemas de sólo recolección de datos a esquemas de análisis de los mismos. En este breve artículo se tratará de brindar una explicación general de algunas metodologías, en este caso serán la metodología Kimball y la metodología Inmon.

Abstract

Data warehouses (data warehouses in English) are becoming increasingly important, as organizations move from data-only schemes to data analysis schemes. In this short article we will try to provide a general explanation of some methodologies, in this case they will be the Kimball methodology and the Inmon methodology.

I. INTRODUCCIÓN

Estamos viviendo en la era de una revolución de datos, y más corporaciones se están dando cuenta de que para liderar, o en algunos casos, para sobrevivir, necesitan aprovechar su riqueza de datos de manera efectiva. El almacén de datos, debido a su propuesta única como repositorio empresarial integrado de datos, está jugando un papel aún más importante en esta situación. Actualmente se practican dos estilos de arquitectura prominentes para construir un almacén de datos: la arquitectura Inmon y la arquitectura Kimball. Este documento intenta comparar y contrastar los pros y los contras de cada estilo de arquitectura y recomendar qué estilo seguir basándose en ciertos factores.

II. OBJETIVOS

A. General:

Dar una visión clara de BI, desde las perspectivas de los autores que sentaron las bases que son Ralph Kimball y Bill Inmon, para la mejora de las estrategias del negocio al que se desee implementar las herramientas de BI.

B. Específicos:

Describir las metodologías propuestas por los principales autores de BI desde las perspectivas de sus creadores Ralph Kimball y Bill Inmon.

III. MARCO TEÓRICO

A. METODOLOGIA SEGUN INMON

Según Bill Inmon, propone que un DW es el conjunto de datos orientados a temas específicos de la organización, es decir deberán ir cambiando con el transcurso del tiempo, pero dichos datos, no deben ser volátiles (no es posible eliminar, ni modificar la información). Al momento de generar un cambio en los datos, se debe de efectuar de una manera que, quede reflejado el cambio que se efectuó y mantener siempre la integridad de la información, se espera que todo resultado del proceso y transformación de la información por medio de herramientas de BI, sea un firme apoyo para los tomadores de decisiones al momento de orientar una estrategia de negocio. .

Consta de las siguientes características:

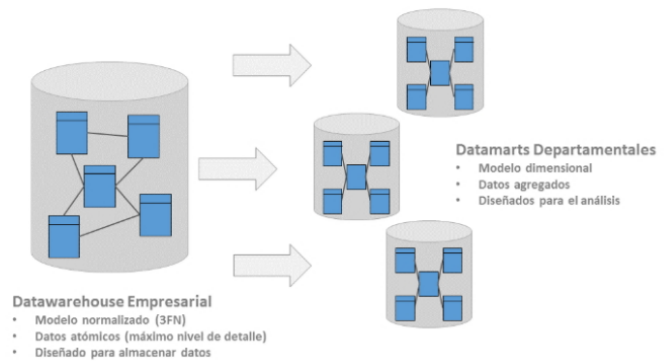
- **Orientado a temas:** Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí.
- **Integrado:** La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes.
- **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.
- **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.



B. ARQUITECTURA SEGUN INMON

La arquitectura que plantea Bill Inmon consta de las siguientes partes: [?]

- **Fuente de la Información:** Inicia el proceso de creación de un DW, conociendo la información que se necesita de todas las herramientas de las que se tengan acceso, ir a las necesidades de información que se necesitan con la finalidad de un resultado para crear un DW.
- **Data Warehouse:** La necesidad de normalizar toda la información extraída para ser almacenada en un DW, los cuales serán procesados y consultados por un DM.
- **Data Marts:** Se crean un subconjunto de los datos de un DW con el objetivo de responder a un determinado análisis o necesidad de una población, de un departamento en específico.
- **Explotación de los datos:** Se refiere a la manera de presentación de la información para ser consultada y analizada por las áreas. En cuanto a la arquitectura interna de un DW, Bill Inmon considera las siguientes características:
 - Normalización: el DW debe ser basado y diseñado, conforme al diseño de las bases de datos transaccionales con las que se esté interactuando.
 - Tercera Forma normal: la prioridad es que el modelo de datos esté construido en TFN con lo cual se tenga mayor relación entre los objetos de la base de datos.



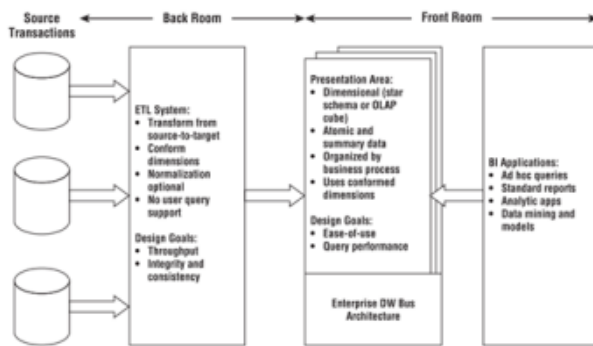
C. METODOLOGIA SEGUN KIMBALL

En su metodología propuesta, menciona la relevancia de la información dentro de las organizaciones y destaca dos propósitos principales: el primero hace referencia a que el sistema operacional que realiza las transacciones de la organización, debe de ser utilizado para el registro de la información y segundo, que los sistemas de DW/BI deben ser utilizados para la consulta de los datos. Según Ralph Kimball, promotor principal del enfoque dimensional define un Data Warehouse como “una copia de los datos transaccionales específicamente para la consulta y el análisis”(6), también parte de su metodología es la forma de construcción del DW. A su vez define que un Data Mart es “Un repositorio de información similar a un DW, pero orientado a una área o departamento”. Propone algunas metas para la buena realización de los sistemas de Negocios Inteligentes:

- Acceso a la información de manera rápida y fácil.
- Credibilidad de la información, los datos deben pasar por un proceso de depuración de la información, descartando la información innecesaria y no relevante, verificación de la fuente de los datos.
- Adaptabilidad a través del tiempo, donde la Información debe estar acompañada del Tiempo.
- Aplicación de controles de seguridad.

La construcción de una solución de DW/BI (Datawarehouse/Business Intelligence) es sumamente compleja, y Kimball nos propone una metodología que nos ayuda a simplificar esa complejidad. Las tareas de esta metodología (ciclo de vida) se muestran en la figura 1.

4.4.1. Arquitectura Ralph Kimball



Para el desarrollo de su esquema los detalla en tres áreas.:

- **Source Transactions:** En esta área está la base transaccional de la organización, es decir la fuente de los datos.
- **Room:** Es donde se encuentra el DW, para su llenado de información hace uso de los sistemas ETL.
- **FrontRoom:** Conformado por el área de presentación y aplicaciones BI es la parte final de la arquitectura donde se organiza y se logra visualizar la información por parte de o de los usuarios finales. plan del proyecto.

ETL. Extracción, transformación y cargado (en inglés extract, transformation, and load), está en el área de back room, entre la fuente de datos y el área de presentación. La extracción es la primera etapa la de obtención de los datos con el fin de leer, entender y copiar los datos requeridos.

4.4.2. Ciclo de vida

Gran parte de la metodología está basado en el ciclo de vida, el cual consta de 4 principios básicos:

- **Proyecto BI centralizado en el negocio:** enfocar todos los esfuerzos a nivel técnico y humano en desarrollar los requerimientos del proyecto con el negocio, dar respuesta por medio de la investigación y el análisis de resultados.
- **Diseñar una infraestructura de datos apropiada:** Diseñar e implementar una base de datos (o de información) única, de fácil uso y alto rendimiento, y este integrada las plataformas de la organización.
- **Realizar entregas:** la carga de los datos en los DW debe ser gradual, dentro de un plazo de 12 meses, esto está condicionado a los requerimientos del negocio.

D. COMPARACION DE METODOLOGIAS

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.[?]

	Inmon	Kimball
Enfoque	Top-Down	Bottom-up
Complejidad de Métodos	Bastante Complejo	Bastante Simple
Estructura de Datos	Métricas comerciales, medidas de desempeño y cuadros de mando	Datos no métricos y datos que se aplicarán para satisfacer múltiples y variadas necesidades de información.
Diseño	Modelado normalizado basado en la empresa	Esquema de Estrella
Objetivos	Ofrezca una solución técnica sólida basada en métodos probados.	Ofrezca una solución que facilite a los usuarios finales la consulta directa de datos y que aún tenga una tasa de respuesta razonable
Herramientas	ERDs y DIS	Parte del modelado relacional tradicional
Principal Audiencia	Profesionales de TI	Usuarios comerciales o finales
Accesibilidad del usuario final	Baja	Alta
Presupuesto	Coste inicial alto	Coste inicial bajo
Plazos	Requiere más tiempo de desarrollo	Tiempo de desarrollo inferior
Requisito de habilidad	Equipo con especialización alta	Equipo con especialización media
Alcance	Toda la compañía	Departamentos individuales
Mantenimiento	Fácil mantenimiento	Mantenimiento más complejo

IV. ANÁLISIS

- Podríamos decir que el enfoque de Kimball se ajusta mas a proyectos pequeños en los que se persiga un sistema fácilmente explotable y entendible por el usuario y de rapido desarrollo, siendo el modelo de Inmon mas apropiado para sistemas complejos de mayor envergadura.

V. CONCLUSIONES

- La metodología de Kimball es ideal para los primeros pasos de implantación de BI a un cliente, cuando la complejidad de almacenamiento de datos no es demasiado grande y donde la infraestructura del BI se encarga de los datos procedentes de un número limitado de fuentes. Sin embargo, cuando el almacén de datos adquiere complejidad, entonces es peligroso forzar el desarrollo de esta metodología. En el mundo del BI, cuando las cosas adquieren gran complejidad, es el momento

de introducir nuevos enfoques al problema, como el propuesto por Inmon.