

Artículo 3. Using OLAP and multidimensional data for decision making

José María García García

11 de marzo de 2016

1. PREGUNTAS

Leer el artículo y obtener de él información acerca de los siguientes puntos.

- **¿Qué dos sistemas puede utilizar un gestor para obtener un mejor conocimiento de su empresa?**

Procesamiento analítico en línea y Bases de Datos de Múltiples Dimensiones. Estos sistemas nos presentan información resumida que han extraído de las bases de la empresa. Cogen datos sobre el rendimiento de la empresa, los trocean y organizan en estructuras de varias dimensiones y los presentan de manera que permite a los gestores descubrir problemas más rápidamente.

- **¿Qué diferencia hay entre los sistemas OLTP y OLAP?**

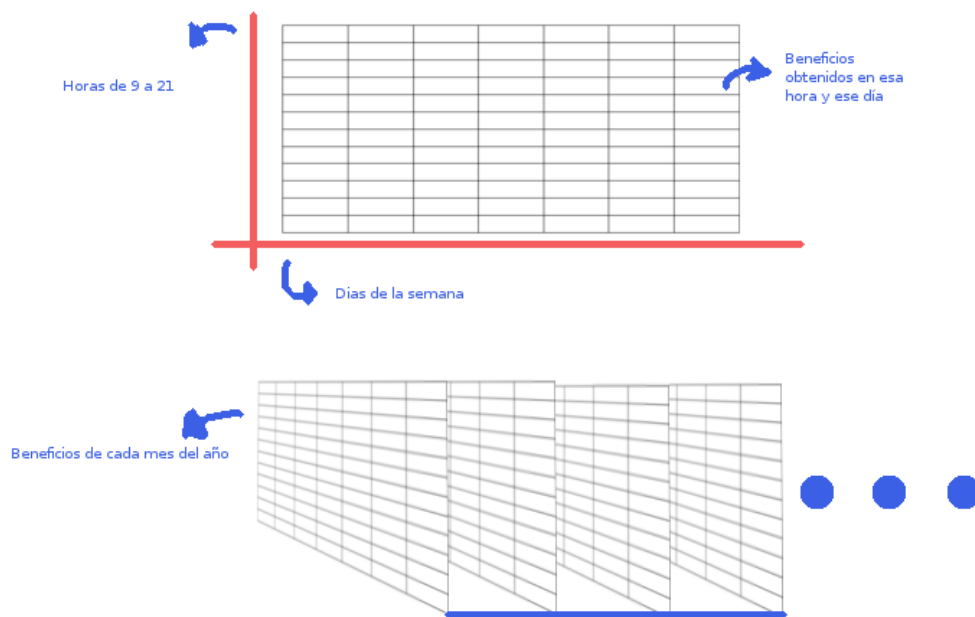
Los sistemas OLTP son sistemas de procesamiento de transacciones en línea, por lo que están centrados en las operaciones del día a día de la empresa, y no tanto en el análisis de los datos que generan esas transacciones. No obstante, estos sistemas suelen incorporar mecanismos para generar informes según las actividades de la empresa, pero estos no sirven para tener una visión general ni para apreciar rápidamente puntos fuertes y puntos débiles.

- **A la hora de obtener información sobre el funcionamiento general de la empresa, ¿qué características tiene el tipo de sistema preferido por los gestores?**

A diferencia del sistema comentado anteriormente, un sistema que usa OLAP y MDDB es intuitivo, fácil de usar y está orientado a la búsqueda de información de tal forma que permita, mediante un vistazo rápido, saber que está pasando en la empresa y formarse una idea general sobre la misma.

- **¿Cómo se representa la información cuando utilizamos una base de datos multidimensional?. Indica un ejemplo**

La información se representa mediante arrays (vectores, listas) de varias dimensiones. A la estructura que toman los datos en la base se la llama **hipercubo**, aunque esta puede no tener forma de cubo, esto es, si el array en cuestión tiene una dimensión 4, no podemos representar esa estructura como un tetraedro, ya que para ello el array debería ser de dimensión 3. Estas estructuras son preferidas a otras porque permiten “jugar” con la información para observarla desde muchos puntos de vista diferentes, centrar el análisis en un único aspecto de los datos, etc. El uso más simple de estas bases es el que podemos observar en el ejemplo.



En este caso, el dato a estudiar ha sido el beneficio que obtiene la empresa y las variables según las cuales lo hemos estudiado han sido la hora del día, el día de la semana y el mes del año. Esta representación permite ver como varían los datos según cada una de las variables, localizar los puntos donde se detecten pérdidas y arreglarlos, etc. Cabe también destacar que cada valor de beneficio está identificado de forma unívoca por sus coordenadas en el cubo (mes, día, hora).

- **¿Por qué las base de datos relacionales son adecuadas para almacenar información transaccional y no analítica?**

Porque estas bases permitían guardar información de las transacciones de forma rápida y sencilla, información que luego se podía estudiar con los informes autogenerados o a través

de órdenes SQL. No obstante, dada la fragmentación y volumen que presentan los datos usados en analítica, manejar estas bases de datos dejaba de ser una trivialidad, por lo que para analítica se prefieren las MDDB.

- **¿Qué tipo de información suele almacenarse en las base de datos multidimensionales?**

Los datos comúnmente almacenados en estas bases de datos suelen ser numéricos y muestran el rendimiento que ha tenido la empresa u organización a lo largo de su historia. Estos datos provienen de otras bases de la empresa, que se vuelcan en la base multidimensional y esta limpia, resume y procesa la información para el uso de los gestores.

- **¿Qué tipo de interactividad es permitida sobre los datos de una base de datos multidimensional? ¿Qué tipo de aplicaciones permiten esta interactividad?**

Siguiendo la metáfora del cubo respecto a la representación de la información, el cubo puede hacerse rodajas, para observar los valores de los datos respecto a un valor concreto de una de las coordenadas de los mismos. Estas rodajas pueden trocearse además para observar la información enmarcada en un rango de valores. Podemos observar cuadrados concretos del cubo o subconjuntos del mismo. Podemos pasar el cubo por un filtro para quedarnos con los bloques que satisfagan ese filtro. Normalmente esto se hace mediante interfaces gráficas, lo que hace el análisis más ameno que la lectura de un informe. Al conjunto de herramientas que se usan para esa manipulación del cubo se les llama OLAP.

- **¿Qué opciones tenemos a la hora de implantar una MDDB? ¿En qué se centra cada una de ellas?**

Existen dos enfoques a la hora de implementar una MDDB:

1. Enfoque top-down (de arriba a abajo): centrado en la planificación del negocio.
2. Enfoque bottom-up (de abajo a arriba): centrado en bases y sistemas existentes.

- **¿En qué consiste el “síndrome de los datos disponibles”?**

Es un síndrome en el que se puede caer al usar como enfoque de implementación el bottom-up, que implica que al partir de datos y sistemas que ya existen, se intenta crear una base que lo contenga todo y lo tenga disponible para todos, cuando realmente eso no es necesario.

- **¿Qué pasos se dan en una implantación top-down?**

En este proceso de implementación se empieza desde arriba, es decir, partimos de lo que la empresa quiere y para que lo quiere. Se lleva a cabo una elicitación de requisitos y el desarrollador prepara un prototipo de MDDB que muestra a la empresa haciendo uso de un software OLAP. Este enfoque asegura que la base que se desarrolle se ajustará a las necesidades de la empresa, lo que no siempre es bueno, ya que la base se desarrollará para un fin concreto, además de que se hará sin tener en cuenta los sistemas sobre los que se apoya, lo que puede dar lugar a problemas si estos no se prevén durante el desarrollo.

- **¿Qué pasos se dan en una implantación bottom-up?**
- **¿Qué enfoque alternativo puede plantearse? ¿En qué consiste?**

2. LISTS

2.1. EXAMPLE OF LIST (3*ITEMIZE)

2.2. EXAMPLE OF LIST (ENUMERATE)

1. First item in a list
2. Second item in a list
3. Third item in a list