

Módulo: Bases de datos

Unidad 1: Almacenamiento de la información

Sesión 1. Conceptos Básicos sobre BD y SGBD.

Descripción:

En la presente sesión vamos a poner en práctica el desarrollo de una base de datos desde el punto de vista empresarial, como un proyecto real que tenemos que desarrollar.

Para ello, vamos a suponer que trabajamos en una importante tienda de moda internacional, y nos han pedido que desarrollemos un proyecto para almacenar toda la información relativa a los productos que han vendido, para que así puedan realizar estudios de mercado para mejorar sus ventas.

La información que quieren almacenar es la siguiente:

- Nombre del producto
- Tipo de producto
- Material del producto
- Coste del producto
- Fecha de la venta
- Tienda de la venta
- País de fabricación

A nosotros, como técnicos dentro del proyecto, se nos pide que hagamos el diseño de almacenamiento de la información para cumplir la petición del cliente.

Criterios de Evaluación:

- RA1_a: Se han analizado los sistemas lógicos de almacenamiento y sus características.
- RA1_b: Se han identificado los distintos tipos de bases de datos según el modelo de datos utilizado.
- RA1_c: Se han identificado los distintos tipos de bases de datos en función de la ubicación de la información.
- RA1_d: Se ha evaluado la utilidad de un sistema gestor de bases de datos.
- RA1_e: Se han clasificado los sistemas gestores de bases de datos.
- RA1_f: Se ha reconocido la función de cada uno de los elementos de un sistema gestor de bases de datos.

- RA1_g: Se ha reconocido la utilidad de las bases de datos distribuidas.
- RA1_h: Se han analizado las políticas de fragmentación de la información.
- RA1_i: Se ha identificado la legislación vigente sobre protección de datos.
- RA1_j: Se han reconocido los conceptos de BigData y de la inteligencia de negocios.

Objetivos:

- Conocer el concepto de base de datos.
- Conocer el concepto de sistema de gestión de base de datos.
- Conocer la estructura y funcionalidad de un sistema de gestión de base de datos.

Recursos:

- Acceso a Internet.
- Software ofimático.

Conceptos a revisar previamente:

- Realizar el estudio de los apartados:
 - Sistema lógico de almacenamiento.
 - Ficheros (planos, indexados, acceso directo, entre otros).
 - Bases de datos.
 - Sistemas gestores de bases de datos: : Funciones, componentes y tipos
 - Bases de datos centralizadas y bases de datos distribuidas.
 - Legislación sobre protección de datos
 - Big Data: introducción, análisis de datos, inteligencia de negocios
- Realizar ejercicios resueltos:
 - “Preguntas”.
 - “Ventajas e Inconvenientes BD frente Sistema Archivos”.
 - “Identificación Sistema”.
 - “Fragmentación en BBDD Distribuidas”.
 - “ANSI-SPARC”
- Visualizar el videoconceptos:
 - “Arquitectura ANSI SPARC.
 - “BD vs. SGBD”.
 - “Componentes de un SGBD”.
 - “Fragmentación”.

Resolución de la práctica:

Como se pide en el enunciado, primero tenemos que hacer un estudio de qué sistema de almacenamiento es el más apropiado para esta base de datos, por lo que surge la primera pregunta:

¿Cómo vamos a guardar la información, en un sistema de ficheros o en una base de datos?

Pero esta pregunta es fácil de responder, porque está claro que dado el volumen de datos a almacenar y el tipo de información, es más conveniente almacenar la información en una base de datos, mejor que en un sistema de ficheros. Porque, ¿cuántos ficheros necesitaríamos para almacenar toda la información? ¿Lo guardamos todo en un mismo fichero, cada campo en un fichero, o cómo? Y según lo almacenemos, ¿cómo accedemos a la información? Tendríamos que abrir los ficheros cada vez que necesitemos hacer una consulta o modificar la información, y después realizar su correspondiente cierre. Además, con todo esto es mucho más complejo guardar la consistencia de la información.

Por todo lo expuestos anteriormente, queda claro que la opción más apropiada para nuestro problema es escoger una base de datos como un tipo de almacenamiento.

Una vez establecido una base de datos como tipo de almacenamiento hardware, el siguiente paso es establecer cómo vamos a tratar la información almacenada. Por lo que surge la siguiente pregunta:

¿Cómo vamos a tratar toda la información almacenada, y la nueva que vayamos a almacenar?

Bien, está claro que el hecho de manejar la información nosotros directamente no la podemos realizar, sería complejo, costoso, tedioso, etc. Tardaríamos horas en realizar acciones demasiado simples, imaginarnos la programación tan costosa para realizar una simple lectura o escritura. Por lo tanto, debemos de buscar una alternativa, y la mejor es la de emplear un Sistema de Gestión de Base de Datos (SGBD).

Además de lo comentado anteriormente, otro motivo por el que seleccionar un SGBD es todas las funciones que nos ofrece de forma integrada, sin tener que estar manejándolas nosotros mismos de forma explícita. Estas funciones son:

- Garantizar la integridad.
- Garantizar las actualizaciones simultáneas.
- Recuperación ante fallos.
- Inclusión de restricciones.

Dentro de los diversos **SGBD**, ahora tenemos que **seleccionar qué tipo sería el más conveniente para nuestro proyecto**. Para ello, primero debemos de tener claro los tipos que hay, y de entre ellos, seleccionar uno. Como hemos visto, existen distintos tipos de SGBD según el criterio en el que nos fijemos, así veámoslo uno por uno.

1. Tipos de SGBD según **el modelo lógico**:

- a. Modelo jerárquico.
- b. Modelo en red.
- c. Modelo relacional.
- d. Modelo orientado a objetos.

En nuestro caso, el *modelo jerárquico* sería un poco complejo de definir, ya que sería complejo definir una jerarquía clara entre los elementos que tenemos que almacenar.

Un *modelo en red* se podría definir, pero sería demasiado complejo, ya que manejar una estructura de red con el volumen de datos que vamos a manejar sería demasiado tedioso.

Un **modelo relacional** sería el más apropiado para nuestro problema, ya que podemos almacenar la información en una tabla, y cada elemento que quieren almacenar (atributo) sería una columna de la tabla, y cada producto sería una fila de esa misma tabla.

Por último, el *modelo orientado a objetos* sería demasiado para el tipo de información que tenemos que almacenar, porque lo que tendríamos recursos innecesarios, que simplemente complicarían su manejo.

2. Tipos de SGBD según el **número de usuarios**:

- a. Monousuario.
- b. Multiusuario.

Como es de suponer, en nuestra base de datos no va a usarla una sola persona, y más sabiendo que la empresa es internacional. En cada país habrá, al menos un empleado que acceda a esa información, y dentro del país, puede que sea por tiendas, etc. Por lo que debemos seleccionar es una **multiusuario**.

3. Tipos de SGBD según su **localización**:

- a. Centralizado.
- b. Distribuido.

Dado el tipo de empresa para el que vamos a trabajar, lo más conveniente es que la base de datos fuera **distribuida**. Una estructura podría ser la siguiente: tendríamos un dispositivo por país en el que se encuentra la tienda. Si esta estructura no es suficiente, podríamos tener un dispositivo por ciudad, o al menos por región, en la que se encuentre la empresa. A pesar de que es la más apropiada para nuestro ejemplo, hay que comentarle al cliente que su inconveniente es que eleva los costes económicos, y en el caso de que no entre en el presupuesto, estudiar la posibilidad de una base de datos centralizada.

Si lo cogiésemos centralizado, el servidor debería de ser lo suficientemente potente para soportar el acceso de todos los empleados a la vez. Además, también tenemos que pensar que si cae el sistema por cualquier motivo, perderíamos toda la información, no es lo mismo que caiga sólo un dispositivo de todos los distribuidos, que se caiga el total.

Una vez que hemos dicho que su localización va a ser distribuida, el siguiente paso es pensar en el tipo de fragmentación, que podía ser horizontal o vertical. Como hemos dicho que tendríamos un dispositivo por país o región por la que se encuentra la empresa, en ese caso lo más conveniente sería hacer una fragmentación horizontal teniendo en cada dispositivo las entradas de los productos que se hayan vendido en esa región. La tabla sería de este tipo:

Tabla producto:

Nombre	Tipo	Material	Coste	Fecha	Tienda venta	País de fabricación

4. Tipos de SGBD según los **lenguajes de consulta**:

- a. SQL.
- b. NoSQL.

Dada la extensión y la facilidad de **SQL**, sería el lenguaje más conveniente para manejar la base de datos.

5. Tipos de SGBD según su coste:

- a. SGBD comerciales.
- b. SGBD gratuitos.

Este punto está claro que depende del presupuesto que nos deje la empresa. Pero más que el presupuesto, habría que fijarse en el soporte que tenga el software, ya que normalmente, aunque sean gratuitos, las licencias de explotación comercial tienen un coste. Por lo tanto, ya que se va a pagar, debemos de tener bien claro que vamos a necesitar un buen soporte técnico detrás, ya que se pueden dar situaciones de caídas o fallos y tienen que ser resueltos correctamente y en la menor brevedad posible.

¿Cuál crees que es el más adecuado?