

# 1. Historia y Evolución de los SGBD (Resumen)

## 1. Introducción

El término bases de datos fue escuchado por primera vez en un simposio celebrado en California en 1963.

En una primera aproximación, se puede decir que una base de datos es un conjunto de información relacionada que se encuentra agrupada o estructurada.

Desde el punto de vista informático, una base de datos es un sistema formado por un conjunto de datos almacenados en discos que permiten el acceso directo a ellos.

Por su parte, un sistema de Gestión de Bases de datos es un tipo de software muy específico dedicado a servir de interfaz entre la base de datos, el usuario y las aplicaciones que la utilizan; es decir, una agrupación de programas que sirven para definir, construir y manipular una base de datos, permitiendo así almacenar y posteriormente acceder a los datos de forma rápida y estructurada.

## 2. Orígenes

Los orígenes de las bases de datos se remontan a la Antigüedad donde ya existían bibliotecas y toda clase de registros, se utilizaban entre otras cosas para recoger información sobre las cosechas y censos. Sin embargo, su búsqueda era lenta y poco eficaz y no se contaba con la ayuda de máquinas que pudiesen reemplazar el trabajo manual.

Las bases de datos se desarrollan a partir de las necesidades de almacenar grandes cantidades de información o datos. El concepto de bases de datos ha estado siempre ligado a la informática y se empleó desde la aparición de las primeras computadoras.

En 1884 Herman Hollerith creó la máquina automática de tarjetas perforadas, se trataba de una máquina tabuladora o censadora, basada en tarjetas perforadas.

En la década de los cincuenta se da origen a las cintas magnéticas, para automatizar la información y hacer respaldos, a través de este mecanismo se comenzó

a automatizar información, con la desventaja de que solo se podía hacer de forma secuencial.

Las aplicaciones informáticas de los años sesenta acostumbraban a darse totalmente por lotes (batch) y estaban pensadas para una tarea muy específica.

Cada aplicación utilizaba ficheros de movimientos para actualizar y/o para consultar uno o dos ficheros maestros o, excepcionalmente, más de dos.

Cada vez que se le quería añadir una aplicación que requiriera el uso de algunos de los datos que ya existían y de otros nuevos, se diseñaba un fichero nuevo con todos los datos necesarios esto provocaba redundancia.

A medida que se fueron introduciendo las líneas de comunicación, los terminales y los discos, se fueron escribiendo programas que permitían a varios usuarios consultar los mismos ficheros on-line y de forma simultánea. Más adelante fue surgiendo la necesidad de hacer las actualizaciones también on-line.

A medida que se integraban las aplicaciones, se tuvieron que interrelacionar sus ficheros y fue necesario eliminar la redundancia.

El nuevo conjunto de ficheros se debía diseñar de modo que estuviesen interrelacionados. El acceso on-line y la utilización eficiente de las interrelaciones exigían estructuras físicas que diesen un acceso rápido, como por ejemplo los índices, las multilistas, etc.

Estos conjuntos de ficheros interrelacionados, con estructuras complejas recibieron al principio el nombre de Data Banks, y después, a inicios de los años setenta, el de Data Bases.

La utilización de estos conjuntos de ficheros por parte de los programas de aplicación era excesivamente compleja, de modo que, especialmente durante la segunda mitad de los años setenta, fue saliendo al mercado software más sofisticado: los **Data Base Management Systems**, que aquí denominamos **Sistemas de Gestión de BD (SGBD)**. Una base de datos es un conjunto estructurado de datos que representa entidades y sus interrelaciones. La representación será única e integrada.

Podríamos agregar que una base de datos de un **SI (Sistema de Información)** es la representación integrada de los conjuntos de entidades instancia correspondientes a las diferentes entidades tipo del SI y de sus interrelaciones.

Esta representación informática (o conjunto estructurado de datos) debe poder ser utilizada de forma compartida por muchos usuarios de distintos tipos.

### 3. Década de los 60.

En los 60 se popularizara el uso de los discos, cosa que fue un adelanto muy efectivo en la época, debido a que a partir de este soporte se podía consultar la información directamente, sin tener que saber la ubicación exacta de los datos.

Nacen las primeras generaciones de bases de datos de red y las bases de datos jerárquicas, ya que era posible guardar estructuras de datos en listas y arboles.

Otro de los principales logros de los años sesenta fue la alianza de IBM y American Airlines para desarrollar **SABRE**, un sistema operativo que manejaba las reservas de vuelos, transacciones e informaciones sobre los pasajeros de la compañía American Airlines.

En esta misma década, se llevo a cabo el desarrollo del IDS desarrollado por Charles Bachman, esto supuso la creación de un nuevo tipo de sistema de bases de datos conocido como modelo en red que permitió la creación de un estándar en los sistemas de bases de datos gracias a la creación de nuevos lenguajes de sistemas de información.

CODASYL (**Conference on Data Systems Languages**) era un consorcio de industrias informáticas que tenían como objetivo la regularización de un lenguaje de programación estándar. Aunque trabajaron en varios lenguajes de programación como COBOL, nunca llegaron a establecer un estándar fijo, proceso que se llevo a cabo por ANSI.

## 4. Década de los 70 – Sistemas Centralizados.

Los primeros SGBD de los años sesenta todavía no se les denominaba así. Estaban orientados a facilitar la utilización de grandes conjuntos de datos en los que las interrelaciones eran complejas.

Estos sistemas trabajaban exclusivamente por lotes (batch).

Al aparecer los terminales de teclado, conectados al ordenador central (**Mainframes**) mediante una línea telefónica, se empiezan a construir grandes aplicaciones on-line transaccionales (**OLTP**).

Para escribir los programas de aplicación se utilizaban lenguajes de alto nivel como Cobol o PL/I, se disponía también de instrucciones y de subrutinas especializadas para tratar las BD que requerían que el programador conociese muchos detalles del diseño físico, y que hacían que la programación fuese muy compleja.

Puesto que los programas estaban relacionados con el nivel físico, se debían modificar continuamente cuando se hacían cambios en el diseño y la organización de la BD. La preocupación básica era maximizar el rendimiento: el tiempo de respuesta y las transacciones por segundo.

Edgar Frank Codd, científico informático inglés, definió el modelo relacional a la par que publicó una serie de reglas para los sistemas de datos relacionales a través de su artículo "**Un modelo relacional de datos para grandes bancos de datos compartidos**".

Este hecho dio paso al nacimiento de la segunda generación de los Sistemas Gestores de Bases de Datos.

Como consecuencia de esto, durante la década de 1970, Lawrence J. Ellison desarrolló el Relational Software System, o lo que es lo mismo, lo que actualmente se conoce como Oracle Corporation, desarrollando así un sistema de gestión de bases de datos relacional con el mismo nombre que dicha compañía.

Finalmente IBM desarrolló unas técnicas para construir un sistema de bases de datos relacionales eficientes, las cuales llamó System R; por otro lado Ingres se desarrolló en la UBC en los años de 1974 a 1977.

Ingres utilizaba un lenguaje de consulta, llamado QUEL, dando pie a la creación de sistemas como Ingres Corporación, MS SQL Server, Sybase, PACE Wang, y Britton Lee-. Por su parte, el Sistema R utilizó el lenguaje de consulta Secuela, el cual ha contribuido al desarrollo de SQL / DS, DB2, Allbase, Oracle y SQL Non-Stop.

Posteriormente en la época de los ochenta también se desarrollará el **SQL (Structured Query Language)** o lo que es lo mismo un lenguaje de consultas o lenguaje declarativo de acceso a bases de datos relacionales.

Cabe destacar que ORACLE es considerado como uno de los sistemas de bases de datos más completos que existen en el mundo, actualmente sufre la competencia del SQL Server de la compañía Microsoft y de la oferta de otros Sistemas Administradores de Bases de Datos Relacionales con licencia libre como es el caso de PostgreSQL, MySQL o Firebird que aparecerían posteriormente en la década de 1990.