ADMINISTRACION DE BASE DE DATOS

1. Perspectiva de la administración de bases de datos

TEMAS

1. Perspectiva de la administración de bases de datos 1.1 Administrador de Base de Datos (DBA)

1.2 Análisis de los manejadores de bases de datos

1.3 Consideraciones para elegir un SGBD

1.4 Nuevas tecnologías y aplicaciones de los sistemas de bases de datos

Competencia

- Desarrollo de software utilizando sistemas de bases de datos.
- Reconocer los alcances y las actividades que deben realizarse como parte del trabajo del ABD.

Los SGBD proporcionan herramientas de alto nivel para gestionar un modelo de datos. Es necesario contar con una o varias personas que conozcan a fondo el SGBD en cuestión y se encarguen de administrarlos correctamente, para eso están los DBA (Data Base Administrators, por sus siglas en inglés, administradores de bases de datos). Comúnmente, en aplicaciones sencillas con pocos requisitos, el mismo programador será el que se encargue de gestionar la base de datos. En organizaciones más grandes es común contar con uno o varios DBA que se encarguen de que la base de datos se mantenga activa las 24 horas del día, de optimizarla y de que los clientes del servidor de bases de datos la utilicen de forma correcta.





En los sistemas de gestión de BBDD actuales existen diferentes categorías de usuarios.

Estas categorías se caracterizan porque cada una de ellas tiene una serie de privilegios o permisos sobre los objetos que forman la BD.

El DBA tiene una gran responsabilidad ya que posee el máximo nivel de privilegios. Será el encargado de crear los usuarios que se conectarán a la BD. En la administración de una BD siempre hay que procurar que haya el menor número de administradores, a ser posible una sola persona.



En los sistemas Oracle las categorías más importantes son:

Cuentas administrativas

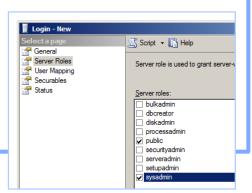
- **SYS.** Inicialmente posee la contraseña CHANGE_ON_INSTALL que, lógicamente, hay que cambiar inmediatamente en la instalación. SYS toma rol de DBA (es decir, de superadministrador) y es en su esquema donde se crea el diccionario de datos; por lo que no conviene de ninguna manera crear otro tipo de elementos en su esquema; es decir, el usuario SYS no debe crear tablas, ni vistas no ningún otro objeto de la base de datos.
- **SYSTEM**. Posee también el rol DBA y se crea durante la instalación. Como antes, la contraseña MANAGER que tiene por defecto se debería cambiar en la instalación. En su esquema se suelen crear tablas y vistas administrativas (pero no se deberían crear otro tipo de tablas).
- **SYSMAN**. Usado para realizar tareas administrativas con la aplicación Database Control del Enterprise Manager.
- **DBSMNP**. Usuario que tiene permisos para monitorizar Enterprise Manager.



SQL Server proporciona roles de nivel de servidor para ayudarle a administrar los permisos de un servidor. Estos roles son entidades de seguridad que agrupan otras entidades de seguridad. Los roles de nivel de servidor se aplican a todo el servidor en lo que respecta a su ámbito de permisos. (Los roles son como los grupos del sistema operativo Windows).

SQL Server además brinda una serie de roles por servidor y por base de datos, que podrán especificarse para cada usuario, aunque también es posible definirlos personalizados. Los roles de servidor son:

- Dbcreator: crea y modifica bases de datos
- Diskadmin: administra los archivos de datos
- Processadmin: administra los procesos de SQL Server
- SecurityAdmin: administra los inicios de sesión.
- Serveradmin: opciones de configuración del servidor
- Setupadmin: instala la replicación
- Sysadmin: realiza cualquier actividad.





En Mysql, es el usuario 'root', el administrador, y que dispone de todos los privilegios disponibles.

Sin embargo, normalmente no será una buena práctica dejar que todos los usuario con acceso al servidor tengan todos los privilegios.

Los conceptos de usuarios y privilegios están íntimamente relacionados. No se pueden crear usuarios sin asignarle al mismo tiempo privilegios. De hecho, la necesidad de crear usuarios está ligada a la necesidad de limitar las acciones que tales usuarios pueden llevar a cabo.

MySQL permite definir diferentes usuarios, y además, asignar a cada uno determinados privilegios en distintos niveles o categorías de ellos.



En MySQL existen cinco niveles distintos de privilegios:

Globales: se aplican al conjunto de todas las bases de datos en un servidor. Es el nivel más alto de privilegio, en el sentido de que su ámbito es el más general.

De base de datos: se refieren a bases de datos individuales, y por extensión, a todos los objetos que contiene cada base de datos.

De tabla: se aplican a tablas individuales, y por lo tanto, a todas las columnas de esas tabla.

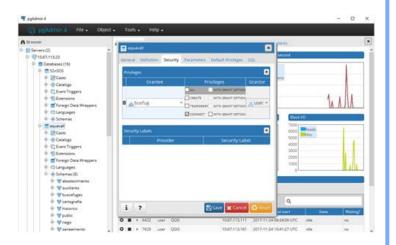
De columna: se aplican a una columna en una tabla concreta.

De rutina: se aplican a los procedimientos almacenados.



En PostgreSQL, los roles son objetos globales que puede acceder a todas las bases de datos de cluster (contando con los privilegios adecuados).

A fin de inicializar un sistema de bases de datos, cada instalación fresca siempre contiene un rol predefinido. Este rol es siempre un **superusuario**, y tiene el mismo nombre (a menos que sea cambiado cuando se corre initdb) que el del usuario (a nivel sistema operativo) que inicializó el cluster de bases de datos. Los roles determinan el conjunto de privilegios disponibles a un cliente conectado.



Es el responsable por la Administración de las Bases de Datos.

Administra las tecnologías de la información y la comunicación, siendo responsable de los aspectos técnicos, tecnológicos, científicos, inteligencia de negocios y legales de bases de datos.

1.1.1 Funciones de un DBA



1.1.1 Funciones de un DBA

• Definición del esquema.

El ABD crea el esquema original de la base de datos escribiendo un conjunto de instrucciones de definición de datos en el LDD.

• Definición de la estructura y del método de acceso.

Establecer estándares de uso, políticas de acceso y protocolos de trabajo diario para los usuarios de la BD. Suministrar la información necesaria sobre la BD a los equipos de análisis y programación de aplicaciones.

• Modificación del esquema y de la organización física.

Los ABD realizan cambios en el esquema y en la organización física para reflejar las necesidades cambiantes de la organización, o para alterar la organización física para mejorar el rendimiento.

Concesión de autorización para el acceso a los datos.

La concesión de diferentes tipos de autorización permite al administrador de la base de datos determinar a qué partes de la base de datos puede acceder cada usuario.

1.1.1 Funciones de un DBA



• Mantenimiento rutinario.

Algunos ejemplos de actividades rutinarias de mantenimiento del administrado

de la base de datos son:

- Copia de seguridad periódica de la base de datos, bien sobre cinta o sobre servidores remotos, para prevenir la pérdida de datos en caso de desastres como inundaciones.
- Asegurarse de que haya suficiente espacio libre en disco para las operaciones normales y aumentar el espacio en disco según sea necesario.
- Supervisión de los trabajos que se ejecuten en la base de datos y asegurarse de que el rendimiento no se degrada por tareas muy costosas iniciadas por algunos usuarios.

1.1.1 Funciones de un DBA

• Efectuar tareas de explotación.

- Vigilar el trabajo diario colaborando en la información y resolución de las dudas de los usuarios de la BD.
- Controlar en tiempo real los accesos, tasas de uso, cargas en los servidores, anomalías, etc.
- Llegado el caso, reorganizar la BD.
- Restaurar la BD después de un incidente material a partir de las copias de seguridad.
- Estudiar las auditorías del sistema para detectar anomalías, intentos de violación de la seguridad, etcétera.
- Ajustar y optimizar la BD mediante el ajuste de sus parámetros, y con ayuda de las herramientas de monitorización y de las estadísticas del sistema.



1.1.2 Relación del DBA con otras áreas de la Sistemas.

Usuarios informáticos. Son los profesionales que definen y preparan la base de datos. Pueden ser:

Directivos/as. Organizadores y coordinadores del proyecto a desarrollar y máximos responsables del mismo. Esto significa que son los encargados de decidir los recursos que se pueden utilizar, planificar el tiempo y las tareas, la atención al usuario y de dirigir las entrevistas y reuniones pertinentes.



1.1.2 Relación del DBA con otras áreas de los Sistemas.

Analistas. Son los encargados de controlar el desarrollo de la base de datos aprobada por la dirección. Normalmente son además los diseñadores de la base de datos (especialmente de los esquemas interno y conceptual) y los directores de la programación de la misma.

Equipo de mantenimiento. Encargados de dar soporte a los usuarios en el trabajo diario (suelen incorporar además tareas administrativas como la creación de copias de seguridad por ejemplo o el arreglo de problemas de red por ejemplo).

Desarrolladores/as o programadores/as. Encargados de la realización de las aplicaciones de usuario de la base de datos.

Administradores/as de las bases de datos. Encargados de crear el esquema interno de la base de datos, que incluye la planificación de copia de seguridad, gestión de usuarios y permisos y creación de los objetos de la base de datos.

1.1.2 Relación del DBA con otras áreas de los Sistemas.



Usuarios de la base de datos

Expertos/as. Utilizan el lenguaje de manipulación de datos (DML) para acceder a la base de datos. Son usuarios que utilizan la base de datos para gestión avanzada de decisiones.

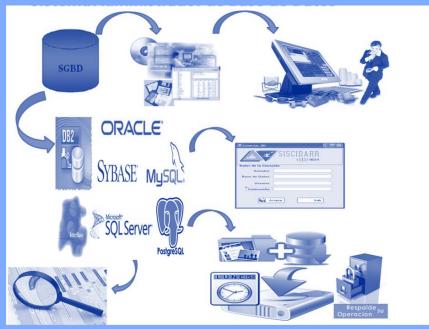
Habituales. Utilizan las aplicaciones creadas por los desarrolladores para consultar y actualizar los datos. Son los que trabajan en la empresa a diario con estas herramientas y el objetivo fundamental de todo el desarrollo de la base de datos.

Ocasionales. Son usuarios que utilizan un acceso mínimo a la base de datos a través de una aplicación que permite consultar ciertos datos. Serían por ejemplo los usuarios que consultan el horario de trenes a través de Internet.

.

1.2. Análisis de los manejad<mark>ores</mark> de bases de datos

Definimos un Sistema Gestor de Bases de Datos o SGBD, como una colección de datos relacionados entre sí, estructurados y organizados, y un conjunto de programas que acceden y gestionan esos datos. La colección de esos datos se denomina Base de Datos.

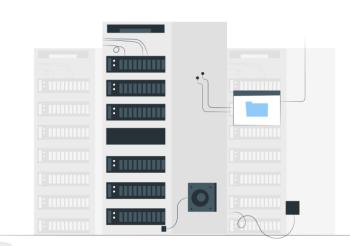


El SGBD es una aplicación que permite a los usuarios definir, crear y mantener la BD y proporciona un acceso controlado a la misma. Debe prestar los siguientes servicios:

- Creación y definición de la BD: especificación de la estructura, el tipo de los datos, las restricciones y relaciones entre ellos mediante lenguajes de definición de datos.
- Toda esta información se almacena en el diccionario de datos, el SGBD proporcionará mecanismos para la gestión del diccionario de datos.
- Manipulación de los datos realizando consultas, inserciones y actualizaciones de los mismos utilizando lenguajes de manipulación de datos.
- Acceso controlado a los datos de la BD mediante mecanismos de seguridad de acceso a los usuarios.
- Mantener la integridad y consistencia de los datos utilizando mecanismos para evitar que los datos sean perjudicados por cambios no autorizados.
- Acceso compartido a la BD, controlando la interacción entre usuarios concurrentes.
- Mecanismos de respaldo y recuperación para restablecer la información en caso de fallos en el sistema.

Algunos de los SGBD más populares actualmente son:

- Microsoft SQL Server
- ✓ Oracle DB
- ✓ MySQL Server





Microsoft SQL Server

- Microsoft SQL Server es un sistema de manejo de bases de datos del modelo relacional, desarrollado por la empresa Microsoft.
- El lenguaje de desarrollo utilizado (por línea de comandos o mediante la interfaz gráfica de Management Studio) es Transact-SQL (TSQL), una implementación del estándar ANSI del lenguaje SQL, utilizado para manipular y recuperar datos (DML), crear tablas y definir relaciones entre ellas (DDL).
- Microsoft SQL Server 2019 Incluye apache Spark y Hadoop Distributed File System (HDFS) para aplicar inteligencia en todos sus datos.
- SQL Server 2022, habilitado por Azure con un rendimiento y una seguridad innovadores. Ahora en versión preliminar privada





Oracle de base de datos

Oracle Database es un sistema de gestión de base de datos de tipo objeto-relacional (ORDBMS, por el acrónimo en inglés de Object-Relational Data Base Management System), desarrollado por Oracle Corporation.

Se considera a Oracle Database como uno de los sistemas de bases de datos más completos, destacando: soporte de transacciones, estabilidad, escalabilidad, y soporte multiplataforma.

Oracle Database ofrece rendimiento, escalabilidad, confiabilidad y seguridad líderes del mercado, tanto en las instalaciones locales como en la nube. Oracle Database 19c es la versión actual a largo plazo, que además proporciona el nivel más alto de estabilidad de la versión y el plazo más largo para asistencia y corrección de errores.

Oracle Database 21c, también disponible para uso en producción hoy como una versión innovadora, proporciona un adelanto de las numerosas mejoras y nuevas funciones. Incluye soporte multimodelo mejorado a través de Javascript en la base de datos y tablas nativas de blockchain, así como mejoras en cargas de trabajo múltiple como AutoML y mejoras de fragmentación que se incluirán en futuras versiones a largo plazo.



MySQL Server

MySQL es un sistema de gestión de bases de datos relacional, multihilo y multiusuario con más de seis millones de instalaciones.1 MySQL AB —desde enero de 2008 una subsidiaria de Sun Microsystems y ésta a su vez de Oracle Corporation desde abril de 2009—desarrolla MySQL como software libre en un esquema de licenciamiento dual.

MySQL Enterprise Edition incluye el conjunto más completo de características avanzadas , herramientas de gestión y apoyo técnico para lograr los más altos niveles de MySQL escalabilidad, seguridad , fiabilidad y tiempo de actividad . Reduce el riesgo , el coste y la complejidad en el desarrollo, despliegue y gestión de aplicaciones MySQL críticos para el negocio .

MySQL es usado por muchos sitios web grandes y populares, como Wikipedia, Google (aunque no para búsquedas), Facebook, Twitter, Flickr, y YouTube.



Postgres

PostgreSQL es un poderoso sistema de base de datos relacional de objetos de código abierto con más de 30 años de desarrollo activo que le ha valido una sólida reputación por su confiabilidad, robustez de funciones y rendimiento.

Se puede encontrar una gran cantidad de información que describe cómo instalar y usar PostgreSQL a través de la documentación oficial. La comunidad de PostgreSQL proporciona muchos lugares útiles para familiarizarse con la tecnología, descubrir cómo funciona y encontrar oportunidades profesionales.

La selección y establecer el entorno de una base de datos requiere de una gran habilidad, conocimiento y consideración. A continuación se describen los principios que intervienen.

Se realiza en primer lugar un análisis de las necesidades del comprador, a continuación se recogen los factores relevantes a tener en cuenta en el proceso de adquisición y, finalmente, se describe cómo deben ser planteadas las especificaciones técnico - funcionales para la elaboración de las Bases de Especificaciones Técnicas, qué normas, estándares y cláusulas tipo pueden ser de aplicación, y cuál es el cuestionario técnico diseñado para normalizar las ofertas y facilitar su evaluación.





Análisis de necesidades

Antes de comenzar el proceso de selección de un SGBD, hay que establecer con la máxima aproximación posible los requisitos que se deben satisfacer mediante su instalación. A continuación se van a detallar varios puntos a tener en cuenta en este análisis.

- ✓ Modelo de datos y arquitectura de implantación específica
- ✓ Rendimiento transaccional exigible
- ✓ Plataforma/s sobre la/s que debe funcionar
- ✓ Tipo de información que se va a tratar
- ✓ Acceso a los datos
- ✓ Integración en el entorno existente
- ✓ Herramientas de administración
- ✓ Características de multiproceso
- ✓ Conectividad y comunicaciones



Modelo de datos y arquitectura de implantación específica

Actualmente, la práctica totalidad de las instalaciones se están realizando con SGBDs relacionales, instalados bien en modo centralizado, bien en modo cliente/servidor y, en pocas ocasiones, arquitecturas distribuidas. Esta es sin duda la solución más conveniente, casi siempre, y la que proporcionará un mayor número de posibilidades de elección.

Dimensionamiento de la BD

El SGBD deberá garantizar que es capaz de manejar el volumen de datos requerido. Para ello, deberá comprobarse que es adecuado en cada uno de los siguientes puntos:

Número total de bases de datos que se van a crear. Número total de tablas por base de datos. Numero máximo de filas por tabla. Longitud máxima de fila. Número máximo de índices por tabla. Número máximo de campos por índice.



Rendimiento transaccional exigible

Si se va a utilizar el SGBD en un entorno transaccional, se deberá conocer cuál es la carga (en transacciones por segundo o por minuto) que deberá soportar el sistema y también, se debe indicar cuál es el tiempo de respuesta aceptable (máximo y medio).

Plataforma/s sobre la/s que debe funcionar

Se deberá especificar la plataforma o plataformas, físicas y lógicas, sobre las que debe funcionar el SGBD. Para cada una se deberá especificar, al menos, el fabricante, modelo y sistema operativo (especificando el número de versión).

Tipo de información que se va a tratar

Todos los productos incluyen soporte para una serie de datos básicos: alfanuméricos, numéricos (enteros y decimales), empaquetados, lógicos y fecha. Según las necesidades específicas de cada caso se deberá exigir el soporte de tipos de datos especiales tales como gráficos, información textual, etc.



Acceso a los datos

Para el acceso a los datos debería evaluarse, además del acceso desde el lenguaje propio del SGBD (si existe), la existencia de herramientas de usuario final tales como generadores de informes, formularios de entrada de datos, etc. También, la posibilidad de acceder desde los lenguajes que se estén utilizando previamente como COBOL, PL/I, etc., que suelen estar soportados por medio de precompiladores.

Integración en el entorno existente

Si la buena integración de cualquier producto informático con el resto de la instalación es siempre una materia de gran importancia, en el caso de un SGBD es absolutamente imprescindible dadas sus características como almacén integrado de datos.

Aunque todo SGBD suele tener alguna herramienta de desarrollo propia, en las instalaciones con desarrollos previos es importante examinar si se podrá acceder a la base de datos con las antiguas herramientas y a qué coste. Este punto determinará, en gran medida, si es posible integrar las aplicaciones antiguas con el SGBD.



Herramientas de administración

El número de las herramientas a disposición del administrador suele ser muy variable según el SGBD. Tener en cuenta que es frecuente que parte de ellas se comercialicen como opciones separadas.

Típicamente existirán herramientas para definir la estructura física, cargar los datos a partir de ficheros secuenciales externos y viceversa, copias de seguridad, utilidades para reorganizar la base de datos para mejorar su eficiencia, aumentar o reducir su tamaño, etc.

También se debe valorar la existencia de herramientas para gestionar la seguridad de monitorización y de obtención de estadísticas.

Características de multiproceso

En entornos donde se requiere un alto rendimiento, puede ser interesante que el SGBD soporte multiproceso. Esto permite que una BD sea accedida por varios procesos que están ejecutándose a la vez, en distintos procesadores y, por tanto, evita las contenciones debidas a sobrecarga del procesador.



Conectividad y comunicaciones

Cuando se necesite poder acceder a BDs situadas en varias máquinas, habrá que asegurarse que el SGBD es capaz de hacerlo o que se incluye el producto adecuado que lo permite. Igualmente, se debe comprobar que funciona con el protocolo de comunicación bajo el que se desea trabajar.



La sofisticación de la tecnología moderna de las bases de datos es el resultado de la evolución que a lo largo de varias décadas ha tenido lugar en el procesamiento de los datos y en la gestión de la información. La tecnología de acceso a los datos se ha desarrollado desde los métodos primitivos de los años cincuenta hasta los potentes e integrados sistemas de hoy en día, arrastrados de un lado por las necesidades y las demandas de la administración de otro, restringida por las limitaciones de la tecnología.

Las nuevas aplicaciones están llevando hasta el límite a los sistemas de bases de datos disponibles, al incorporar documentos multimedia. Imágenes, series temporales, datos activos, grandes cantidades de información (no olvidemos que los datos se expanden hasta llenar el espacio disponible), etc.

Los usuarios cada vez necesitamos más recursos en tecnología, es por eso que surgen las evoluciones de sistemas, y por en de las bases de datos.



La expectativa del futuro de las bases de datos se proyecta a la manipulación de bases de datos por medio de todos dispositivos móviles, con acceso móvil a la red, esto empieza en nuestra actualidad así como la seguridad de las bases de datos para aspectos militares, médicos, entre otros, además de las bases de datos paralelas en las cual se maneja la misma base de datos al mismo tiempo, las bases de datos en memoria principal la cual se ubica en la memoria principal lo que elimina o al menos disminuye la necesidad de operaciones de entrada/salida, las bases de datos activas las cuales se producen cuando ciertas condiciones se ejecutan de manera automática, sin intervención del usuario, las bases de datos deductivas, las bases de datos orientadas a objetos, las bases de datos multimedia, las bases de datos temporales y las bases de datos difusas, entre otras son en la actualidad lo que estamos viviendo en tecnología de almacenamiento de datos, así que esperemos a ver si todo esto tiene sus frutos en los próximos 10 años.

En los últimos años venimos asistiendo a un avance espectacular en la tecnología de bases de datos. Temas que hasta hace poco parecían exclusivos de laboratorios y centros de investigación, comienzan a aparecer en las ultimas versiones de algunos SGBD y en nuevos productos: bases de datos multimedia, activas, deductivas, orientadas a objetos, seguras, temporales, móviles, paralelas, difusas, etc.

NoSQL ("Not Only SQL")

Las bases de datos NoSQL están diseñadas específicamente para modelos de datos específicos y tienen esquemas flexibles para crear aplicaciones modernas. Las bases de datos NoSQL son ampliamente reconocidas porque son fáciles de desarrollar, por su funcionalidad y el rendimiento a escala.

Ej. Cassandra, Redis y MongoDB



Apache Cassandra se trata de un software NoSQL distribuido y basado en un modelo de almacenamiento de «clave-valor», de código abierto que está escrita en Java. Permite grandes volúmenes de datos en forma distribuida

NewSQL

NewSQL es una clase de sistemas modernos de gestión de bases de datos relacionales que tratan de conseguir el mismo rendimiento escalable de sistemas NoSQL para el procesamiento de transacciones en línea, manteniendo durante las cargas de trabajo las garantías ACID de un sistema de base de datos tradicional

Ej. CockroachDB, Google Cloud Spanner, Clustrix, Azure SQL, y NuoDB

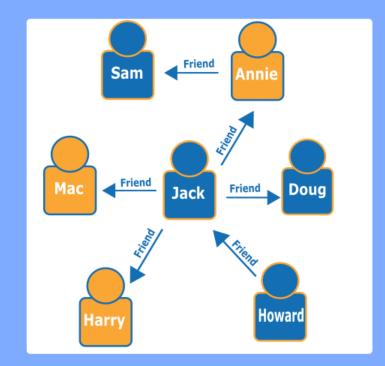


Azure SQL es una familia de servicios de base de datos SQL seguros y totalmente administrados que se basan en el mismo motor de SQL Server

Bases de datos graficas

Las bases de datos gráficas funcionan almacenando las relaciones junto con los datos. Debido a que los nodos relacionados están físicamente vinculados en la base de datos, el acceso a esas relaciones es tan inmediato como el acceso a los datos en sí, simplemente leen la relación desde el almacenamiento. Las consultas satisfactorias son una simple cuestión de caminar o "atravesar" el gráfico.

Ej. Neo4J, OrientDB, y DataStax



Ejemplo de un gráfico de red social. Dadas las personas (nodos) y sus relaciones (bordes), puede averiguar quiénes son los "amigos de los amigos" de una persona en particular, por ejemplo, los amigos de los amigos de Sebastián.

Bases de datos en la nube

Una base de datos en la nube es un servicio de base de datos creado y accedido por medio de una plataforma en la nube. Sirve a muchas de las mismas funciones que una base de datos tradicional, con la flexibilidad añadida de la computación en la nube.

Ej. Amazon Aurora , <u>Google Cloud</u> <u>Database</u>



Amazon Aurora es una base de datos relacional compatible con MySQL y PostgreSQL creada para la nube.