



PRÁCTICA 1

TEMA 1

C.E.S ACADEMIA LOPE DE VEGA

CFGS: 2º Administración de Sistemas Informáticos en Red

Curso: 2017/2018

Asignatura: ADMINISTRACIÓN DE SISTEMAS GESTORES DE BASES DE DATOS

Prof. Miguel Castro Ortiz

Autor: Rafael Osuna Ventura

1. Averigua y explica el significado del término ACID en el contexto de los sistemas gestores de bases de datos.

ACID son las siglas de Atomicidad, Consistencia, Aislamiento y Durabilidad. Estas son un conjunto de propiedades necesarias para que un conjunto de instrucciones, sean consideradas como una transacción en un sistema de gestión de bases de datos, es decir, que las ordenes se ejecuten de una forma indivisible o atómica.

- Atomicidad: Significa que el sistema permite operaciones atómicas. Deben ejecutarse todas correctamente, o en el caso de que alguna de ellas no pueda hacerlo, el efecto de las que ya se han ejecutado no debe hacerse notar, debe deshacerse, como si los conjuntos de las operaciones no se hubieran realizado.
La atomicidad está íntimamente ligada al concepto de transacción de los sistemas gestores de bases de datos.
- Consistencia: Integridad, esta propiedad asegura que sólo se empieza aquello que se puede acabar. Por lo tanto, se ejecutan aquellas operaciones que no van a romper las reglas y directrices de integridad de la base de datos. Sostiene que cualquier transacción llevará a la base de datos desde un estado válido a otro también válido.
- Aislamiento: Propiedad que asegura que una operación no puede afectar a otras. Esto asegura que la realización de dos transacciones sobre la misma información sea independiente y no generen ningún tipo de error.
- Durabilidad: Propiedad que asegura que, una vez realizada la operación, ésta persistirá y no se podrá deshacer, aunque falle el sistema.

2. Averigua en qué consiste y para qué sirve la minería de datos.

La minería de datos se define como la extracción no trivial de información implícita, previamente desconocida y potencialmente útil a partir de datos, es decir, la exploración y el análisis de grandes cantidades de datos con el fin de descubrir patrones con significado.

La minería de datos comprende cinco elementos principales:

1. Procesos ETL: extraer, transformar y cargar datos transaccionales en los sistemas de almacenamiento de datos.
2. Almacén de datos: recopilar y gestionar los datos en un sistema de base de datos multidimensional.
3. Acceso a los datos: proporcionar acceso a los datos a los analistas de negocio, profesionales de tecnología de información y usuarios autorizados.
4. Análisis de datos: analizar los datos de todas las aplicaciones.
5. Presentación: de los datos en un formato útil, como un gráfico o una tabla, que facilite una buena interpretación de su contenido.

3. Investiga que base de datos creada por defecto utiliza MySQL para almacenar el diccionario de datos. De todas las tablas que forma parte del diccionario de datos. Selecciona las 5 que creas que ofrece información muy relevante. Explica la información que almacena y lista los campos que componen cada una de las 5 tablas.

Por defecto usa el fichero SCHEMA .El soporte para INFORMATION_SCHEMA está disponible en MySQL 5.0.2 y posterior. Proporciona acceso a los metadatos de la base de datos.

4. Asocia cada función que realiza un SGBD con el componente encargado de llevarla a cabo.

Operaciones básicas → DDL,DCL,DML,TCL

Catálogo → Diccionario de Datos

Garantizar la integridad → Integridad, transacciones

Permitir actualizaciones → Control de concurrencia, se usa DML

Recuperación de datos → Control de recuperación, importar/exportar datos, replicación, copias de seguridad

Integración → programas de aplicaciones

Cumplir restricciones → objetos, tablas base, seguridad

Herramientas de administración → Importar/exportar datos, sincronización, optimizador de consultas, planificador

5. Define DDL, DCL, DML y TCL. A continuación, para cada uno de los lenguajes escribe las sentencias SQL que están relacionadas. Pon un ejemplo de uso de cada sentencia.

Las sentencias DDL se utilizan para crear y modificar la estructura de las tablas, así como otros objetos de la base de datos.

- CREATE - para crear objetos en la base de datos.
- ALTER - modifica la estructura de la base de datos.
- DROP - borra objetos de la base de datos.
- TRUNCATE - elimina todos los registros de la tabla, incluyendo todos los espacios asignados a los registros.

Las sentencias de lenguaje de manipulación de datos (DML) son utilizadas para gestionar datos dentro de los schemas. Algunos ejemplos:

- SELECT - para obtener datos de una base de datos.
- INSERT - para insertar datos a una tabla.
- UPDATE - para modificar datos existentes dentro de una tabla.
- DELETE - elimina todos los registros de la tabla; no borra los espacios asignados a los registros.

Es un lenguaje de programación y un subconjunto de SQL, que se utiliza para controlar el procesamiento de transacciones en una base de datos.

- COMMIT - guarda el trabajo realizado.
- SAVEPOINT - identifica un punto en una transacción a la que más tarde se puede volver.
- ROLLBACK - restaurar la base de datos a la original, hasta el último commit.
- SET TRANSACTION - cambia las opciones de transacción como nivel de aislamiento y qué segmento de cancelación utiliza.

6. ¿Qué es integridad de los datos? Explica cada uno de los tipos de restricciones de integridad que existen. Pon un ejemplo que aclare el concepto de cada tipo.

Integridad de datos se refiere al estado de corrección y completitud de los datos ingresados en una base de datos. Hay varios tipos:

-Integridad de entidad: define una fila como entidad única para una tabla determinada. La integridad de entidad exige la integridad de las columnas de los identificadores o la clave principal de una tabla, mediante índices y restricciones UNIQUE, o restricciones PRIMARY KEY.

-Integridad de dominio: viene dada por la validez de las entradas para una columna determinada. Puede exigir la integridad de dominio para restringir el tipo mediante tipos de datos, el formato mediante reglas y restricciones CHECK, o el intervalo de valores posibles mediante restricciones FOREIGN KEY, restricciones CHECK, definiciones DEFAULT, definiciones NOT NULL y reglas.

-Integridad referencial: protege las relaciones definidas entre las tablas cuando se crean o se eliminan filas. En SQL Server la integridad referencial se basa en las relaciones entre claves externas y claves principales o entre claves externas y claves exclusivas, mediante restricciones FOREIGN KEY y CHECK. La integridad referencial garantiza que los valores de clave sean coherentes en las distintas tablas. Para conseguir esa coherencia, es preciso que no haya referencias a valores inexistentes y que, si cambia el valor de una clave, todas las referencias a ella se cambien en consecuencia en toda la base de datos.

-Integridad definida por el usuario: permite definir reglas de empresa específicas que no pertenecen a ninguna otra categoría de integridad. Todas las categorías de integridad admiten la integridad definida por el usuario. Esto incluye todas las restricciones de nivel de columna y nivel de tabla en CREATE TABLE, procedimientos almacenados y desencadenadores.

7. Atendiendo al modelo lógico en el que se basan los distintos SGBD podemos diferenciar varios tipos. Explica cada uno de ellos.

- Relacionales: En este modelo todos los datos son almacenados en relaciones, y como cada relación es un conjunto de datos, el orden en el que estos se almacenen no tiene relevancia (a diferencia de otros modelos como el jerárquico y el de red). Esto tiene la considerable ventaja de que es más fácil de entender y de utilizar por un usuario no experto. La información puede ser recuperada o almacenada por medio de consultas que ofrecen una amplia flexibilidad y poder para administrar la información.

- En Red: Los datos en el modelo de red se representan mediante colecciones de registros y las relaciones entre los datos se representan mediante enlaces, que se pueden ver como punteros. Los registros en la base de datos se organizan como colecciones de grafos dirigidos.
- Jerárquicos: es similar al modelo de redes, en el sentido en que los datos y las relaciones entre los datos se representan mediante registros y enlaces, respectivamente. Éste se diferencia del modelo de redes en que los registros se organizan como colecciones de árboles en lugar de grafos dirigidos.
- Orientados a objetos: está basado en una colección de objetos. Un objeto contiene valores almacenados en variables de ejemplares dentro de ese objeto. Un objeto también contiene fragmentos de código que operan en el objeto. Estos fragmentos de código se llaman métodos. Los objetos que contienen los mismos tipos de valores y los mismos métodos se agrupan juntos en clases. Una clase se puede ver como una definición de tipo para los objetos. Esta combinación de datos y métodos constituyendo una definición de tipo es similar a un tipo abstracto de datos en un lenguaje de programación

8. Averigua el significado de DNS y LDAP. ¿Con qué base de datos está relacionada (relacional, jerárquica o en red)? Justifica tu respuesta.

DNS: Este protocolo se utiliza para poder recordar de manera sencilla las direcciones IP. De esta manera surge el concepto de nombres de dominio. Gracias a esto podemos asignar a una dirección IP un nombre, además de que es más fiable porque la dirección IP de un servidor puede cambiar, pero el nombre no lo hace. Podemos decir entonces que el DNS es un sistema jerárquico y distribuido que permite traducir nombres de dominio en direcciones IP y viceversa.

LDAP: hace referencia a un protocolo a nivel de aplicación que permite el acceso a un servicio de directorio ordenado y distribuido para buscar diversa información en un entorno de red. LDAP también se considera una base de datos (aunque su sistema de almacenamiento puede ser diferente) a la que pueden realizarse consultas.

Está relacionada con la base de datos jerárquica ya que cada entrada posee un conjunto de atributos que se organizan en forma de árbol y cada nodo debe poseer un nombre único.

9. Averigua el significado de los siguientes acrónimos en el contexto de los SGBD (SAP, ERP y DSS).

SAP: Una base de datos lógica (BDL) es un programa SAP especial abap/4 que lee y procesa información de tablas de una base de datos. Puede incluirse como atributo de un programa abap/4 de usuario, y se compila conjuntamente con él.

Desde la transacción SE36 podemos crear o administrar base de datos lógica.

ERP: Enterprise Resource Planning. Aplicaciones informáticas (software), que permiten gestionar e integrar a una empresa que además de la producción, controlan los aspectos financieros, logísticos (por ejemplo, ventas) y de recursos humanos de manera integrada, tanto

con referencia a los datos como a los procedimientos operativos. Afecta a toda la empresa y controla los recursos necesarios para la gestión integral de la misma.

DSS: generalmente se compone de reglas y mecanismos, una base de conocimiento organizacional donde se encuentran diferentes alternativas a la solución de un problema específico. Con esto, el administrador puede visualizar que pasará si decide tomar una decisión o si decide cambiarla y combinarla con otros escenarios.

10. Haz un listado de al menos 3 SGBD libres (open source) y 3 comerciales indicando 3 de sus características principales. Discute los motivos por los que consideras que algunas empresas ofrecen productos software gratuitos.

Libres:

-MySQL:

- Contiene un sólido y amplio subconjunto del lenguaje SQL.
- Dispone una amplia cantidad de plataformas transacciones y claves foráneas.
- Contiene una conectividad segura y una búsqueda e indexación de campos de texto.
- Integración perfecta con PHP

-ApacheDerby:

- El tamaño del programa es muy reducido, ocupando 2.6 megabytes entre el motor base y controlador integrado JDBC, utilizando muy poco espacio en el computador donde se instalará el programa.
- Es un software libre.

-DB2:

- Ofrece a los desarrolladores un conjunto de herramientas para crear aplicaciones de bases de datos para el escritorio, entornos cliente/servidor, tablet PC o para la Web.

Comerciales:

-Oracle:

- Es una herramienta de administración gráfica que es muy intuitiva y cómoda de utilizar.
- Ayuda a analizar datos y efectuar recomendaciones para mejorar el rendimiento y la eficiencia en el manejo de aquellos datos que se encuentran almacenados.
- También apoya en el diseño y optimización de modelos de datos

- Microsoft SQL Server:

- Posee una compresión de Datos que permite que los datos se almacenen de una manera más eficiente, y reduzca los requisitos de almacenamiento para sus datos.
- La Compresión de Datos también ofrece mejoras significativas en el rendimiento para grandes cargas de trabajo

-FoxPro:

- Ofrece a los desarrolladores un conjunto de herramientas para crear aplicaciones de bases de datos para el escritorio, entornos cliente/servidor, tablet PC o para la Web.
- Entre sus características se pueden enumerar:
- Capacidades poderosas y muy veloces para el manejo de datos nativos y remotos.
- Flexibilidad para crear todo tipo de soluciones de bases de datos.

Algunas empresas venden software gratuito ya que su estrategia no se basa en vender el producto sino un servicio, es decir, en dar soporte de calidad a ese producto y así generar un producto de mejor calidad que sus competidores siendo este gratis.

11. Comenta ventajas e inconvenientes de las bases de datos centralizadas respecto a las bases de datos distribuidas.

Centralizadas:

- Control centralizado: un solo administrador de la base de datos
- Independencia de datos: la organización de los datos es transparente para el programador
- Reducción de redundancia: una sola copia de datos que se comparte
- Estructuras físicas complejas para accesos eficientes
- Seguridad

Distribuidas:

- Control jerárquico: existe un administrador global y otro local
- Transparencia en la Distribución: la localización de los datos es un aspecto adicional de la independencia de datos
- Replicación de datos: copias múltiples de datos que incrementa la localidad y la disponibilidad de datos
- Uso de optimización global para reducir la transferencia de datos
- Problemas de seguridad intrínsecos

12. Para garantizar la integridad, el SGBD puede utilizar varios mecanismos, uno de ellos es el bloqueo. La siguiente tabla muestra de forma gráfica el problema que podría haber si no controlamos dicha concurrencia. Completa las siguientes tablas y observa la diferencia entre no utilizar control de concurrencia (tabla1) y controlar la concurrencia realizando bloqueos (tabla 2).

P: READ A; A:=A+1; WRITE A;

A en la BD	5	5	5	5	6	6
T1	READ A		A:=A+1			WRITE A
T2		READ A		A:=A+1	WRITE A	
A en T1	5		6			6
A en T2		5		6	6	

P: LOCK A; READ A; A:=A+1; WRITE A; UNLOCK A;

A en la BD		5	5	6	6	6	6	7	7
T1	LOCK A	READ A	A:=A+1	WRITE A	UNLOCK A				
T2			LOCK A	LOCK A	LOCK A	READ A	A:=A+1	WRITE A	UNLOCK A
A en T1		5	6	6					
A en T2						6	7	7	

13. Investiga los posibles mecanismos que tiene un SGBD para realizar un control de recuperación de una base de datos ante un fallo crítico.

Un mecanismo es tener siempre una copia de seguridad para así tener todos los datos almacenados de una forma segura en caso de pérdida. Otro mecanismo es crear checkpoints para así en caso de fallo volver a un punto anterior de la base de datos y no perder todos los datos.