**Preguntas y respuestas teóricas – Parcial**

**¿Qué son los archivos de montículo? Describa brevemente sus características.**

Hay varias organizaciones primarias de archivos que determinan la forma en que los registros de un archivo se colocan físicamente en el disco. Los archivos de montículo (o archivos no ordenados) colocan los registros en disco sin un orden específico, en tanto que los archivos ordenados (o archivos secuenciales) mantienen los registros ordenados según el valor de un cierto campo. Los archivos dispersos utilizan una función de dispersión para determinar la colocación de los registros en el disco. Otras organizaciones primarias de los archivos, como los árboles B, se valen de estructuras de árbol.

En el tipo más simple y básico de organización (montículo) , los registros se colocan en el archivo en el orden en que se insertan, y los registros nuevos se insertan al final del archivo. Un archivo con este tipo de organización se denomina archivo de montículo o de montón. Esta organización suele utilizarse con caminos de acceso adicionales, como los índices secundarios.

La inserción de un registro nuevo es muy eficiente: el último bloque del archivo en disco se copia en un buffer, se añade el nuevo registro, y se reescribe el bloque en el disco.

Para eliminar un registro, el programa encargado de ello primero tendrá que hallarlo, copiar el bloque en un buffer, eliminar el registro del buffer y finalmente reescribir el bloque en el disco. Esto deja espacio desocupado adicional en el bloque, así que la eliminación de un gran número de registros de esta forma origina una gran cantidad de espacio desperdiciado.

**Dentro de las técnicas de dispersión se encuentra la “Dispersión Interna”. Explique el concepto.**

Otro tipo de organización primaria de archivos se basa en la dispersión (hashing), que proporciona un acceso muy rápido a los registros con ciertas condiciones de búsqueda. Esta organización suele recibir el nombre de archivo disperso o directo. La condición de búsqueda debe ser una condición de igualdad sobre un solo campo, el campo de dispersión del archivo. Es común que el campo de dispersión sea también un campo clave del archivo, en cuyo caso se habla de la clave de dispersión. La dispersión se basa en establecer una función h, llamada función de dispersión o función de aleatorización, que se aplica al valor del campo de dispersión de un archivo y produce la dirección del bloque de disco en el que está almacenado el registro. La búsqueda del registro dentro del bloque puede realizarse en un buffer de la memoria principal. Para la mayoría de los registros, basta un solo acceso a bloque para obtener un registro. La dispersión también sirve como estructura interna de datos en un programa siempre que se tenga acceso a un archivo temporal pequeño empleando el valor de un campo.

**De todas las propiedades que poseen las relaciones, cuales tienen que ver con el cuerpo de la relación. Explique cada una de ellas justificando su respuesta.**

Las relaciones poseen ciertas propiedades, todas ellas consecuencias inmediatas de la definición de relación. Las propiedades dentro de cualquier relación dada son:

* No existen tupias duplicadas:

Esta propiedad surge del hecho de que el cuerpo de la relación es un conjunto matemático (de tuplas) y en consecuencia no incluyen elementos duplicados. Esta primera propiedad sirve para ilustrar el punto de que una relación y una tabla no son lo mismo, ya que una tabla puede contener filas duplicadas, mientras que una relación no puede contener una tupla duplicada. (Ya que si una "relación" contiene tupias duplicadas, entonces, por definición, ¡no es una relación!)

* Las tupias están en desorden, de arriba hacia abajo:

Esta propiedad también surge del hecho de que el cuerpo de la relación es un conjunto matemático y por definición, los conjuntos no están ordenados. Esta segunda propiedad sirve también para ilustrar la idea de que una relación y una tabla no son lo mismo, ya que las filas de una tabla tienen un ordenamiento (de arriba hacia abajo o viceversa), mientras que las tupias de una relación no lo tienen.

* Los atributos están en desorden, de izquierda a derecha:

Esta propiedad surge del hecho de que el encabezado de una relación también es un conjunto (de atributos). Por lo tanto, no existe algo como "el primer atributo" o "el segundo atributo" (etcétera) y tampoco hay "el siguiente atributo" (de nuevo, no existe el concepto de "el siguiente"), en otras palabras, siempre se hace referencia a los atributos por nombre, nunca por posición.

* Cada tupla contiene exactamente un valor para cada atributo:

Esta propiedad surge inmediatamente de la definición de una tupia: una tupia es un conjunto de n componentes o pares ordenados. Se dice que una relación que satisface esta cuarta propiedad está normalizada o, lo que es equivalente, está en la primera forma normal. Esta propiedad en particular podría parecer muy obvia y de hecho lo es, en especial debido a que todas las relaciones están normalizadas de acuerdo con la definición.

**¿Qué son los archivos ISAM? ¿Describa brevemente su funcionamiento? ¿Es una estructura fácil de migrar? ¿por qué?**

ISAM son siglas de Indexed Sequential Access Method (Método de Acceso Secuencial Indexado),

* Método para almacenar información a la que se pueda acceder rápidamente.
* La información se organiza en registros compuestos por campos de tamaño fijo.
* Los registros se almacenan secuencialmente, inicialmente para acelerar el acceso en sistemas de cinta.
* Un conjunto secundario de ficheros dispersos (tablas "hash") conocidos como índices contienen «punteros» a los registros que permiten acceder a los registros individuales sin tener que buscar en todo el fichero.
* Este es el punto de partida para todos los modernos sistemas de bases de datos navegacionales, en los cuales los punteros que dirigen hacia otra información fueron almacenados dentro de los propios registros.
* El avance clave que posee ISAM es que los índices son pequeños y pueden ser buscados rápidamente, permitiendo a la base de datos acceder sólo a los registros que necesita.
* Modificaciones adicionales a la información no requieren cambios a otra información, sólo a la tabla y los índices.

Las bases de datos relacionales pueden fácilmente ser construidas en una red ISAM con la adición de lógica para mantener la validez de los enlaces entre las tablas. Típicamente el campo usado como enlace, será indexado para su lectura rápida. ISAM es muy fácil de entender e implementar, puesto que básicamente consiste en el acceso directo y secuencial a un archivo en una base de datos.

**Explique qué es y para qué sirve la propiedad de cierre de algebra relacional. ¿Qué relación tiene dicha propiedad sobre las cabeceras de las relaciones?**

Propiedad de cierre: Esta me garantiza que cualquier operación que yo realice mediante álgebra, me va a devolver algo que va a tener todas las características de una relación con todas sus propiedades.

En cuanto a la cabecera de las relaciones, la propiedad de cierre me garantiza que tanto los atributos de la cabecera como el dominio de cada uno de ellos mantendrán sus características.

**¿Cuáles son las técnicas de dispersión dinámica? Describa el funcionamiento de cada una.**

En la dispersión dinámica el número de cubetas no es fijo, como en la dispersión normal estática.

Las técnicas de dispersión dinámica son: Dispersión DINAMICA, dispersión EXTENSIBLE y dispersión LINEAL.

* Dinámica: El archivo comienza con una cubeta, si se desborda se divide en 2. Bit 0 camino izquierdo bit 1 camino derecho.
* Extensible: Se basa en un arreglo de 2^d donde d es la profundidad global. Si d=3 entonces bloque de 3 bits.
* Lineal: Aumenta o disminuye el archivo de dispersión sin necesidad de un directorio. Se van dividiendo las cubetas en orden lineal a medida que se provoque el desborde.

**Dispersión Interna y Externa. Explicar conceptos, ventajas, desventajas.**

La dispersión es una técnica primaria de organización de archivos. Se utiliza el campo clave de dispersión.

* Interna: Se utiliza para dispersión en archivos internos y se implementa con un arreglo de registros. La desventaja de este método es que produce colisiones, las cuales son controladas mediante una tabla de dispersión (ocupada entre el 70% y el 90%).
* Externa: Se basa en el almacenamiento de archivos en disco. Se emplea mediante cubetas. La ventaja de este método es que no se originan tantos problemas de colisiones como en la dispersión interna, ya que aquí se aplica el método de cubetas mediante desborde.

**¿Cuáles son los tipos de conjuntos presentes en el modelo de red? Indique cuales son las propiedades básicas. ¿Y cuáles son los tipos especiales de conjuntos? Descríbalos.**

Un tipo de conjunto es una descripción de un vinculo 1:N entre dos tipos de registros.

Las propiedades básicas de un tipo de conjunto son: un NOMBRE para el tipo de conjunto, un PROPIETARIO y un MIEMBRO.

Los tipos especiales de conjuntos son: propiedad del SISTEMA (no tiene registro propietario el propietario es el sistema), MULTIMIEMBRO (los conjuntos miembro de un registro pueden pertenecer a más de un tipo de registro) y RECURSIVOS (el mismo tipo de registro desempeña tanto un papel de propietario como de miembro).

**Describa el funcionamiento de los árboles B y B+ por los índices. ¡Qué diferencia hay entre ellos? ¿Porque suele preferirse un arbol B+ como estructura de acceso a los archivos de datos.**

Los arboles B y B+ poseen índices de múltiples niveles.

Los árboles b y b+ son estructuras de acceso de datos con forma de árbol que pueden ser usadas en bases de datos. La diferencia entre los árboles b y b+ es que el acceso en los arboles b+ es más rápido ya que todos los nodos de una misma altura se encuentran enlazados en forma de lista. Además los arboles B+ permiten lograr índices con menos niveles y más alta capacidad.

**Diferencias entre administrados de BD y administrador de Datos.**

El administrador de datos es el encargado de tomar las decisiones y las políticas de lo que se va a hacer con los datos (tiene una visión global del negocio). El administrador de base de datos es el que implementa las políticas definidas por el administrador de datos, otorgando permisos, creando estructuras, entre otras.

La principal diferencia del DBA y el DA es que el DA define Políticas y el DBA implementa Políticas.

El DA no es un usuario de la base de datos, ya que cumple otras funciones.

**Defina independencia lógica y física. Beneficios**

* Independencia física: Permite que se almacenen los datos independientemente de la forma lógica en que serán presentados.
* Independencia lógica: Permite que las aplicaciones puedan obtener la información de diferentes formas, sin importar como se encuentren almacenados. La principal ventaja es que es posible modificar la parte física de una base de datos sin que cambien su estructura lógica y viceversa.

**Características sistemas de sección frontal y sección posterior. Enumerar sistemas de sección frontal.**

* De la sección frontal del DBMS: Son un utilitarios, sirven para cumplir alguna funcionalidad, algún propósito. Ej: Hacer un listado. Todos estos utilitarios interactúan con el DBMS. Están destinados al usuario final y eventualmente a un programador.
  + Utilitario tipo Reporte/Gráficos.
  + Utilitarios generadores de códigos.
  + Herramientas CASE.
* De la sección posterior del DBMS: Son del tipo herramientas, sirven para reparar algo, para evaluar o controlar algo. Ej: Ver performance. Los utilitarios de esta sección no siempre interactúan con el DBMS o por lo menos no de manera constante (puede hacer una consulta inicial y luego trabajar solos). Son de uso exclusivo del DBA.

**Archivos ordenados y no ordenados**

Existen 3 métodos de organización de archivos en una base de datos: ordenados, no ordenados y dispersos.

* No ordenados: Es el tipo más simple y básico de organización de archivos. Los registros se colocan en el archivo en el orden que se insertan y los nuevos se insertan al final del archivo. Se utilizan índices secundarios como camino de acceso adicional. Insertar un registro es muy eficiente, pero buscar un determinado archivo es muy costoso ya que se busca bloque por bloque. Desperdicia mucho espacio al borrar.

Beneficio: Inserción y eliminación rápida.

Contra: La búsqueda es costosa.

* Ordenados: Los archivos ordenados, se ordenan en el disco en base a un campo de ordenación. El campo de ordenación es también llamado campo clave cuando el campo por el que se ordena identifica unívocamente a ese archivo. La lectura de estos archivos es muy eficiente y la búsqueda de algún registro mediante búsqueda binaria genera una velocidad muy superior que otros métodos. No es muy eficiente si se busca por un campo que no sea el campo clave de ordenamiento. Insertar o eliminar archivos es muy costoso ya que hay que mantener el orden.

Beneficio: El método de acceso y la búsqueda de un registro es muy rápida.

Contra: La inserción y eliminación es muy costosa.

**¿Qué es y para qué sirve la clave foránea?**

Una clave foránea es un mecanismo de INTEGRIDAD que se establece entre dos relaciones (no necesariamente distintas), donde los valores de un subconjunto de atributos de la cabecera de la primera relación, deben coincidir con los valores de la clave primaria de la segunda relación. A la primera relación se la denomina relación referencial y a la segunda relación referida. Y estas claves foráneas deben cumplir las siguientes dos reglas, independientemente del tiempo.

**¿Por que resulta costoso tener acceso a un bloque de disco?**

Por los tiempos que se consumen con: Mover cabezal, esperar posición, leer datos.

**Analice las distintas técnicas para eliminar registros.**

La eliminación de un registro se da en archivos ordenados y no ordenados.

* En los no ordenados, debe encontrarlo, copiar el bloque a un buffer, eliminarlo y reescribir el bloque en el disco. Otra forma es tener un byte de marcador de eliminación.
* En los ordenados, es una operación costosa porque se debe conservar el orden. Se puede utilizar marcadores de borrado y actualizarlo periódicamente.

**¿Qué diferencia hay entre índices primarios, índices secundarios e índices de agrupamiento?**

Hay varios tipos de índices de un solo nivel, entre ellos los primarios, secundarios y de agrupamiento. Están compuestos por 2 campos, uno de indexación y una dirección de bloque de disco/registro según sea denso o no.

* Índice primario: Archivo ordenado de longitud fija, con 2 campos. El primero es del mismo tipo de dato que el campo clave del archivo de datos, el segundo un apuntador al bloque de disco. El campo clave se denomina Clave primaria. Debe tener el mismo ordenamiento que los datos. Es no denso, hay una entrada por cada bloque de disco del fichero de datos, y no una entrada por cada registro. Sobre un fichero ordenado por clave solo puede definirse un índice primario.
* Índice de agrupamiento: Cuando los registros están ordenados según un campo no clave y no único dicho campo se denomina, campo de agrupamiento. Un índice de agrupamiento es también un registro agrupado por dos campos: el primero es del mismo tipo del campo clave y el segundo un apuntador a un bloque. Hay una entrada en el índice como valores diferentes del campo clave. Índice no denso. Sobre un fichero ordenado por un campo no clave (campo de agrupamiento) sólo se puede definir un índice de agrupamiento.
* Índices secundarios: Archivos ordenados con 2 campos. El primero corresponde a algún tipo no de ordenamiento del archivo, y se denomina campo de indización. El segundo apunta a un bloque o a un registro. Pueden existir varios. Es denso.



Si el campo del índice secundario tiene varias ocurrencias del mismo se implementa un nivel intermedio, llamado bloque de apuntadores de registro, en donde se sigue la siguiente lógica:



Propiedades de los índices:



**Diferencia entre índice denso y no denso**

* Índice denso: Si el índice tiene una entrada por cada registro del archivo se denomina denso.
* Índice no denso: El índice tiene una entrada que apunta al principio de una página del archivo (que puede estar guardada en disco o en memoria), donde dentro de la misma se encuentra guardado el registro.

**¿Cómo mejora la eficiencia de búsquedas en un archivo de índices con la indización de múltiples niveles?**

Se pretende reducir el trozo del índice sobre el que seguir buscando. Normalmente tienen 3 niveles. Las hojas le apuntan directamente al registro en cuestión. Por eso tienen 2 campos en su estructura. Cuando se llenan se parten a la mitad. Este esquema multinivel se puede utilizar sobre cualquier tipo de índice, sea primario, de agrupamiento o secundario, siempre que el primer nivel tenga valores distintos en el campo de indexación y entradas de tamaño fijo. Cuando el índice del primer nivel es denso, se puede saber si existe un registro con un valor dado en el campo de indexación (test de existencia) sin tener que acceder al fichero de datos, cosa que hay que hacer siempre en un índice no denso. Por lo tanto los índices multinivel reducen el número de accesos a bloque cuando se busca un registro a partir del valor de su campo de indexación.



**¿Qué diferencia hay entre una organización de archivo y un método de acceso?**

Organización de archivo:

* Registros de longitud fija.
* Registros de longitud variable: La longitud de los registros debe estar definida correctamente para poder leer y escribir de forma efectiva.

Método de acceso:

* Acceso secuencial: Archivo secuenciales y de texto.
* Acceso Directo: Permite recuperar registros individuales sin leer otros registros del archivo, ejemplos archivos indexados.

**¿Qué razones hay para tener registros de longitud variable?**

Puede ser variable por:

* Los registros son todos del mismo tipo pero hay campos con tamaño variable.
* Los registros son todos del mismo tipo, pero hay campos con múltiples valores en registros individuales.
* Los registros son todos del mismo tipo, pero uno o más de los campos son opcionales.
* El archivo contiene registros de diferentes tipos por lo que su tamaño es variable.

**¿Cuáles son los aspectos de los que se encarga el modelo relacional?**

Éste es el modelo más utilizado en la actualidad para modelar problemas reales y administrar datos dinámicamente. Está definido en base a un modelo matemático de conjuntos. Su idea fundamental es el uso de "relaciones". Estas relaciones podrían considerarse en forma lógica como conjuntos de datos. Esto es pensando en cada relación como si fuese una tabla que está compuesta por registros(las filas de una tabla), y campos(las columnas de una tabla). En este modelo, el lugar y la forma en que se almacenen los datos no tienen 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 para un usuario esporádico de la base de datos. La información puede ser recuperada o almacenada mediante "consultas" que ofrecen una amplia flexibilidad y poder para administrar la información. El lenguaje más habitual para construir las consultas a bases de datos relacionales es SQL, Structured Query Language, un estándar implementado por los principales motores o sistemas de gestión de bases de datos relacionales.

* Percepción: De manera tabular, bidimensional.
* Ubicación: El usuario no tiene porque saber donde están los datos (que trabaja con el modelo lógico).
* Lenguaje: Lenguaje natural.

Los grandes proveedores de modelos de RED se ajustaron creando una capa de soft para parecerse al relacional. Comenzó como un modelo teórico y luego como un producto.

**¿Cuáles son los operadores tradicionales del álgebra relacional? Defina a cada uno de ellos junto con sus propiedades.**

Hay 3 grupos de operadores:

* Operaciones tradicionales de conjuntos o básicos: Intersección, unión, diferencia, producto y división.
* Operaciones relaciones especiales: Restricción, reunión y proyección.
* Operaciones adicionales: Extender, resumir y renombrar.

**¿Cuáles son las principales funciones del DBMS?**

EL DBMS (sistema de administración de base de datos) es el que maneja los accesos a la base de datos. Como ejemplo:

1. Un usuario emite una petición de acceso en SQL.
2. El DBMS intercepta la petición y la analiza.
3. El DBMS inspecciona el esquema externo del usuario, la transformación externa/conceptual, el esquema conceptual, la transformación interna/conceptual y la definición de estructura de almacenamiento.
4. El DBMS ejecuta las acciones sobre la base de datos almacenada.

Funciones del DBMS:

* Definición de datos: Debe ser capaz de aceptar definiciones de datos en la forma fuente y convertirlas a la forma fuente correspondiente. Contiene un compilador y un procesador DDL para cada nivel.
* Manipulación de datos: Debe ser capaz de manejar peticiones de recuperar actualizar o eliminar datos. Debe tener un procesador DML o compilador DML para tratar con el DML.
* Las peticiones planeadas y no planeada: La planeada esta contemplada desde el diseño físico de la BDD y la no planeada puede no estarlo.
* Optimización y ejecución
* Seguridad e integridad de los datos: Debe vigilar al usuario por infracciones de seguridad.
* Recuperación de datos y ocurrencias.
* Diccionario de datos: El diccionario contiene metadatos. En general todos los diversos esquemas y transformaciones, y restricciones de seguridad están almacenadas en el DDD.
* Rendimiento: Debe realizar las tareas de la manera mas eficiente.

Podemos definir entonces que la finalidad del DBMS consiste en proporcionar una interfaz para el usuario con la BD.

**¿Cuáles son los distintos tipos de usuario existentes? ¿El administrador de datos es uno de ellos? ¿Cuál es su función?**

Existen 3 grandes grupos:

* Programadores: Realizan los programas y se comunican realizando las llamadas al DBMS correspondiente. Les interesan las estructuras.
* Usuarios Finales: Interactúan con la base desde terminales. A partir del modelo de Red comenzaron a aparecer como usuarios finales las personas.
* Administrador de Base de Datos (DBA): Crea las bases e implementa los controles técnicos. Es el encargado de que el sistema funcione con el rendimiento adecuado.

Estos 3 son los que definen la base de datos. Dentro de los perfiles, existe uno que es el administrador de datos.

Administración de datos: El administrador de datos, decide que datos se guardan en las bases, y establece políticas para manejarlos (seguridad de datos por ejemplo). Actualmente el administrador de datos quedó en desuso. Es alguien que conoce a la organización.

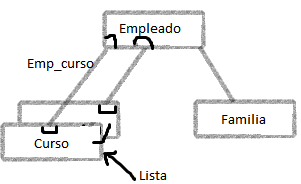
Junto con los problemas de seguridad los administradores de datos, surgieron junto con el área de seguridad informática.

**Defina el modelo red.**

Hay dos estructuras de datos básicos, registros y conjuntos (relación/vínculo 1:N entre 2 tipos de registros). Los registros se clasifican por tipo de registro, le damos un nombre a cada tipo y a cada atributo. Se creó para solventar problemas en el modelo jerárquico.

Si lo comparamos con el modelo jerárquico, en vez de ser un árbol, se puede ingresar por cualquier registro que se defina. No necesariamente el más alto de la jerarquía del árbol. Se cambia el concepto de Padre-hijo, hacia el concepto de Dueño-miembro. Puede haber un registro con más de un Dueño o padre. Cada registro tiene su propia cadena de apuntadores.

Un ejemplo sencillo de esto sería así:



En este esquema el empleado puede tener muchos cursos. Esta estructura me permite tener todas las cardinalidades posibles, por ejemplo la cardinalidad N:M pueden ser 2 sets. El mantenimiento de esta estructura es bastante engorroso. La navegación se produce registro a registro. Al ir navegando registros a registro no se puede tener toda la estructura en memoria, sino que se debe ir guardando la información parcializada en archivos. Por los problemas que generaban las cardinalidades posibles es que surgen la normalización o modelado de datos. El modelo de red sigue un estándar, una norma que define como debe comportarse con accesos, registros, etc.

**Dentro de las técnicas de dispersión se encuentra la “Dispersión Interna”. Explique el concepto.**

Son técnicas de Hash, algún tipo de algoritmo que tenga alguna distribución. La distribución interna coloca los desbordes en el mismo archivo que está manejando. Trata de distribuir los registros uniformemente en el espacio de direcciones de modo que haya el mínimo de colisiones y a la vez se ocupen la mayor parte de las posiciones.

**Dentro de las técnicas de dispersión se encuentra la “Dispersión Externa”. Explique el concepto.**

Son técnicas de Hash, algún tipo de algoritmo que tenga alguna distribución. La distribución externa coloca los desbordes en bloques o páginas de datos para manejar los desbordes. Tiene una clave, que es relativo al número de bloque. Hay una tabla que se maneja en el descriptor del archivo, convierte las claves en la dirección del bloque de disco. Ofrece acceso más rápido para obtener un registro arbitrario en un archivo. Reduce las colisiones.

**Defina los conceptos de, base de datos y datos persistentes.**

* Base de datos: Una BD es el conjunto de componentes de datos, software y hardware que permiten administrar, persistir y explotar datos de una empresa (persiga esta objetivos con o sin fines de lucro).
* Datos persistentes: Un dato pasa a ser persistente cuando queda bajo el control de la BD, sin importar cuanto tiempo ese dato permanezca en la BD. Esto es, el usuario (sea una persona o un sistema usuario) ingresa un dato mediante los medios que ofrece la BD para tal fin (puede ser vía un lenguaje o una utilidad) y a partir de ahí pasa a ser un dato persistente, aun en el caso que ese mismo dato sea borrado un instante después por el usuario vía los mecanismos adecuados.

**Defina la operación de producto cartesiano extendido. Como funciona, Que devuelve, Cuales son sus condiciones.**

Dadas las relaciones R de grado m y S de grado n, el producto cartesiano extendido lo que hace es vincular cada elemento de R con uno de S y da como resultado una Relación de grado m+n.

**Cuáles son los datos persistentes de una base de datos.**

Los datos persistentes de una base de datos son exclusivamente aquellos ingresados por usuarios o por algún utilitario de carga. Los datos que definen la estructura de la BD no son datos persistentes.

**¿Qué representa el nivel conceptual?**

LOS TRES NIVELES DE LA ARQUITECTURA

La arquitectura ANSI/SPARC se divide en tres niveles, conocidos como interno, conceptual y externo, respectivamente.

* El nivel interno (también conocido como el nivel físico): Es el que está más cerca del almacenamiento físico; es decir, es el que tiene que ver con la forma en que los datos están almacenados físicamente.
* El nivel externo (también conocido como el nivel lógico de usuario): Es el más próximo a los usuarios; es decir, el que tiene que ver con la forma en que los usuarios individuales ven los datos.
* El nivel conceptual (también conocido como el nivel lógico de la comunidad, o en ocasiones sólo como el nivel lógico, sin calificar): Es un nivel de indirección entre los otros dos. Observe que el nivel externo tiene que ver con las percepciones de usuarios individuales, mientras que el nivel conceptual tiene que ver con la percepción de una comunidad de usuarios.

**¿Cuáles son las principales funciones del DBA (Administrador de Base de Datos)?**

* Definicion del esquema conceptual: Es trabajo del administrador de datos decidir exactamente qué información contendrá la base de datos; en otras palabras, identificar las entidades de interés para la empresa e identificar la información que hay que registrar acerca de dichas entidades. Por lo regular a este proceso se le conoce como diseño lógico(en ocasiones conceptual) de la base de datos.
* Definicion del esquema interno: El DBA también debe decidir la forma en que van a ser representados los datos en la base de datos almacenada. A este proceso se le conoce comúnmente como diseño físico de la base de datos.
* Establecer disponibilidad con los usuarios: Es asunto del DBA enlazarse con los usuarios para asegurar que los datos necesarios estén disponibles y para escribir (o ayudar a escribir) los esquemas externos necesarios, utilizando el DDL externo aplicable.
* Definir restricciones de seguridad e integridad: Como ya expliqué, las restricciones de seguridad y de integridad pueden ser vistas como parte del esquema conceptual. El DDL conceptual debe incluir facilidades para especificar dichas restricciones
* Definir políticas de backup y restauración: Una vez que una empresa se compromete con un sistema de base de datos, se vuelve drásticamente dependiente del funcionamiento exitoso de dicho sistema. En el caso de que se produzca un daño en cualquier parte de la base de datos (ocasionado, por ejemplo, por un error humano o por una falla en el hardware o en el sistema operativo) resulta esencial poder reparar los datos afectados con el mínimo de demora y con tan poco efecto como sea posible sobre el resto del sistema.
* Supervision de rendimiento: Organizar el sistema de tal manera que se obtenga el rendimiento "ideal para la empresa" y de hacer los ajustes apropiados (es decir, afinar) conforme las necesidades cambien.

**Defina el modelo jerárquico**

Son base de datos que almacenan su información en una estructura jerárquica. En este modelo los datos se organizan en una forma similar a un árbol (visto al revés), en donde un nodo padre de información puede tener varios hijos. El nodo que no tiene padre es llamado raíz y los nodos que no tienen hijos se los conoce como hojas. Las bases de datos jerárquicos son especialmente útiles en aplicaciones que manejan un gran volumen de información y datos muy compartidos permitiendo crear estructuras estables y de gran rendimiento. Una de las principales limitaciones de este modelo es su incapacidad de representar eficientemente la redundancia de datos.