

Universidad de Costa Rica

Mario Quirós Luna

B76090

Tarea 3

Datos y la administración del almacenamiento.

La tecnología de almacenamiento que predomina para la gestión de datos es el disco duro, los discos duros modernos son más confiables que los años pasados con el incremento de tiempo entre fallos.

Los mecanismos de estos los hacen más vulnerables a fallos que otros componentes de computadora, cuando el número de discos duros incrementa en un sistema la vulnerabilidad del sistema también incrementa.

El rendimiento de la base de datos depende de la entrada y salida.

El almacenamiento se está convirtiendo en el negocio central de operación, debido a que cada vez se almacenan más y más datos.

Archivos y conjuntos de datos.

Cuando se determina la cantidad de espacio requerida el DBA debe elegir un dispositivo de almacenamiento apropiado con suficiente espacio para la base de datos y todos sus archivos.

Colocación de archivos en el disco

El DBA debe determinar el lugar óptimo para los archivos en los discos, algunas veces el DBA puede ganar rendimiento simplemente moviendo archivos de un disco físico a otro.

Una técnica para optimizar los discos es solo usar la memoria exterior del disco así el brazo no necesita moverse muy lejos para encontrar los datos.

Particiones sin formato contra sistemas de archivos

Una partición sin formato es la preferida para los dispositivos físicos cuando se quiere desarrollar una base de datos en servidores de bases de datos basados en UNIX. Con una partición sin formato es usada, los datos se escribirán físicamente en el disco sin una operación de sistema interviniendo.

El inconveniente con los dispositivos sin formato es la dificultad para rastrear los archivos de la base de datos.

Gestión de espacio

Lidiar con falta de espacio es común y molesto para los DBA, es mejor estar preparado para que esto no ocurra, tirando objetos de bases de datos que ya no se utilizan.

Diseño de páginas de datos

Cada DBMS usa diferente formato para el diseño de las páginas de la base de datos, típicamente el diseño de la página para la base de datos consiste en tres componentes básicos.

- El encabezado de la página.
- Las filas de datos.
- La tabla de compresión.

Páginas de asignación

Las páginas de asignación controlan las paginas físicas, cada página física esta mapeada en una sola base de datos. Las paginas físicas no necesitan estar contiguas al disco.

Diseño de registro de datos

Una tabla está compuesta de filas y columnas, cada fila debe estar enteramente almacenada en una página de datos y está contenida con un registro que consiste en datos adicionales de mantenimiento junto a los datos reales. Cada tabla de registro está compuesta de los siguientes elementos.

- Fila de cabecera.
- Fila de datos.
- Tabla de compresión.

Calculando el tamaño de la tabla

Una vez el diseño de la página está decidido el DBA puede calcular el tamaño del registro, calculando el tamaño de la tabla fácilmente. Después de calcular el tamaño del registro, el siguiente paso es determinar cuántas filas tendrá una página física, el tamaño de la página diferirá entre DBMSs.

Diseño de páginas de índice

Cuando las tablas no tienen índices definidas las filas simplemente se agregan a lo último de la tabla, los registros de índices están compuestos por varias cosas:

- Información de encabezado.
- El tamaño de la fila
- Los valores de los índices.
- Página puntera.
- Tabla de compresión.

Calculando el tamaño del índice

Primeramente, se necesita calcular el tamaño de la fila, se necesita calcular el tamaño del registro de índice, no solo el tamaño de la fila. Una vez se calcula el tamaño del registro, se necesita calcular la cantidad de espacio de almacenamiento requerida por el índice.

Registro de transacciones

Uno de los más importantes archivos para la base de datos es el procesado del registro de transacciones.

La actividad de la base de datos es el factor más importante a la hora de determinar el tamaño apropiado para el registro de transacciones.

Fragmentación y almacenamiento

La fragmentación consiste en separar un archivo en partes a través del disco, hay aspectos de la fragmentación que el DBA debe monitorear y corregir. La fragmentación del disco puede impactar en el rendimiento de algunos sistemas.

Opciones de almacenamiento

Una vez se calcula el tamaño del espacio requerido para los archivos de la base de datos, se necesita escoger un tipo de almacenamiento para utilizar, hay muchas opciones de almacenamiento disponibles que los DBA pueden escoger. Los más comunes son los discos duros, pero también hay otras opciones disponibles, como los dispositivos de estado sólido (SSDs). Los SSD almacenan datos de forma similar a los discos duros, pero a diferencia que el SSD usa memoria en lugar de discos magnéticos el rendimiento de entrada y salida mejora enormemente.

RAIDS

(Resumido en la Tarea 2)

Solo un montón de discos (JBOD)

Es un término utilizado para diferencia las tecnologías de disco tradicionales de la tecnología de almacenamiento más reciente, por lo general, los discos están conectados directamente a un servidor. Una vez que el DBMS está instalado en esas unidades de disco, los archivos de base de datos pueden crear en esas unidades de disco usando el DBMS.

Red de área de almacenamiento (SAN)

Generalmente es una red interconectada de dispositivos de almacenamientos. El SAN provee los siguientes beneficios.

- Almacenamiento compartido por múltiples anfitriones
- Alto rendimiento de entrada y salida
- Consolidación de servidores y almacenamiento

Almacenamiento conectado a la red (NAS)

Se refiere al almacenamiento al que se puede acceder directamente desde la red. Con NAS, los anfitriones o los sistemas cliente pueden leer y escribir datos a través de una interfaz de red utilizando protocolos. El NAS provee las siguientes ventajas

- Almacenamiento compartido entre varios anfitriones
- Gestión más sencilla debido a la reducción del almacenamiento duplicado
- Acceso al almacenamiento basado en aplicaciones a nivel de archivo

SAN vs NAS

Una SAN se utiliza mejor como una red principal de almacenamiento que proporciona servicios de almacenamiento básicos. También, son adecuadas para compartir el almacenamiento y construir la infraestructura para la consolidación del almacenamiento y el servidor. SAN es ideal para aplicaciones de bases de datos.

NAS es más adecuado para resolver problemas de almacenamiento multimedia, problemas de intercambio de datos y almacenamiento compartido, NAS no maneja de manera eficiente el almacenamiento basado en bloques que utilizan los sistemas de bases de datos.

Almacenamiento por niveles

Con el almacenamiento por niveles, se asignan diferentes categorías de datos a diferentes tipos de medios de almacenamiento para reducir el costo total de almacenamiento. Consiste en una técnica de categorización de datos de temperatura múltiple. lo cual divide los datos en 4 categorías: caliente, tibio, frío e inactivo.

Planificación de capacidad

El objetivo es ajustar los recursos disponibles en el sistema según corresponda. Al medir la capacidad actual, medir el crecimiento de la capacidad a lo largo del tiempo y tener en cuenta los requisitos de capacidad anticipados de las nuevas iniciativas corporativas y de TI, puede determinar si su infraestructura existente puede soportar la carga de trabajo anticipada.

Un DBA debe ser capaz de responder estas preguntas:

¿Cuándo se requerirá más almacenamiento? ¿Cuánto almacenamiento adicional se necesita?
¿Dónde se necesita el almacenamiento adicional? ¿Qué se debe hacer para alinear el almacenamiento adicional con el DBMS?

Las bases de datos se componen de datos y archivos que deben almacenarse.

Bibliografía:

Mullins, C. (2002). *Database Administration: The Complete Guide to Practices and Procedures*.
Adison Wesley.
<https://ptgmedia.pearsoncmg.com/images/9780321822949/samplepages/0321822943.pdf>