Optimización y Consulta a través de Índices

Introducción

La optimización de consultas es esencial para mejorar el rendimiento en bases de datos. Los índices juegan un papel crucial en esta optimización, permitiendo búsquedas más rápidas y eficientes.

Creación de una Copia de la Tabla sin Índices

Para demostrar la importancia de los índices, primero crearemos una copia de la tabla 'gasto' sin índices.

-- Crear una copia de la tabla 'gasto' en una nueva tabla 'gasto\_sin\_indices'

SELECT \*

INTO gasto\_sin\_indices

FROM gasto;

Comparación de Consultas

A continuación, compararemos el rendimiento de una consulta en la tabla original con índices y en la nueva tabla sin índices.

Consulta en la Tabla Original

SET STATISTICS TIME ON;

SET STATISTICS IO ON;

SELECT

\* FROM gasto WHERE fechapago BETWEEN '2023-01-01' AND '2023-02-28';

SET STATISTICS TIME OFF;

SET STATISTICS IO OFF;

\*\*\*\*\* Analsis de resultados\*\*\*\*\*

SQL Server Execution Times:

CPU time = 125 ms, elapsed time = 2899 ms.

SET STATISTICS TIME ON;

SET STATISTICS IO ON;

SELECT \*

FROM gasto\_sin\_indices

WHERE fechapago BETWEEN '2023-01-01' AND '2023-02-28';

SET STATISTICS TIME OFF;

SET STATISTICS IO OFF;

\*\*\*\*\*consulta con indice\*\*\*\*\*

SET STATISTICS TIME ON;

SET STATISTICS IO ON;

SELECT

\* FROM gasto WHERE fechapago BETWEEN '2023-01-01' AND '2023-02-28';

SET STATISTICS TIME OFF;

SET STATISTICS IO OFF;

Análisis de Resultados

CPU time: Mide el tiempo de CPU que se utilizó para ejecutar la consulta.

Sin Índice: 125 ms

Con Índice: 281 ms

Elapsed time: Mide el tiempo total transcurrido desde que se inició la consulta hasta que finalizó.

Sin Índice: 2899 ms

Con Índice: 2619 ms

Interpretación

Sin Índice:

El menor tiempo de CPU (125 ms) sugiere que el servidor no necesita procesar índices. Sin embargo, el tiempo total transcurrido (2899 ms) es mayor porque la consulta requiere un escaneo completo de la tabla, lo que consume más tiempo debido a la cantidad de datos.

Con Índice:

El mayor tiempo de CPU (281 ms) indica que el servidor está utilizando el índice para buscar los datos de manera más eficiente. Aunque el tiempo de CPU es mayor, el tiempo total transcurrido (2619 ms) es menor, lo que muestra que el uso del índice mejora el rendimiento general de la consulta al reducir el tiempo necesario para acceder a los datos.

La consulta con índice es más rápida porque reduce la necesidad de escanear toda la tabla, lo que se refleja en el menor tiempo total transcurrido.

Sí, los resultados obtenidos están en línea con lo que se esperaría al comparar consultas con y sin índices: el uso de índices reduce el tiempo total necesario para ejecutar la consulta, a pesar de un incremento en el tiempo de CPU debido a la búsqueda optimizada.

\*\*\*\*\*\*\*\*\*SE COMPARA LOS TIEMPOS DE ACCESO CON INDICE AGRUPADO Y SIN AGRUPAR\*\*\*\*\*\*\*\*

SET STATISTICS TIME ON;

SET STATISTICS IO ON;

-- Ejecutar la consulta en la tabla original con el índice agrupado

SELECT \*

FROM gasto

WHERE fechapago BETWEEN '2023-01-01' AND '2023-02-28';

SET STATISTICS TIME OFF;

SET STATISTICS IO OFF;

\*\*\* su resultado\*\*\*

SQL Server Execution Times:

CPU time = 250 ms, elapsed time = 2881 ms.

\*\*\* indice no agrupado\*\*\*\*

SET STATISTICS TIME ON;

SET STATISTICS IO ON;

-- Ejecutar la consulta en la tabla con el índice no agrupado

SELECT \*

FROM gasto\_sin\_indices

WHERE fechapago BETWEEN '2023-01-01' AND '2023-02-28';

SET STATISTICS TIME OFF;

SET STATISTICS IO OFF;

Análisis de Resultados

CPU time:

Índice Agrupado: 250 ms

Índice No Agrupado: 110 ms

Elapsed time:

Índice Agrupado: 2881 ms

Índice No Agrupado: 2635 ms

Interpretación

Índice Agrupado

CPU time: 250 ms indica que el servidor está utilizando el índice agrupado para buscar los datos. Este mayor uso de CPU es debido a que el índice agrupado ordena físicamente los datos en la tabla, facilitando la búsqueda dentro de un rango de valores.

Elapsed time: 2881 ms sugiere que, aunque el índice agrupado requiere más CPU, ofrece un rendimiento similar en términos de tiempo total transcurrido al buscar datos en un rango específico.

Índice No Agrupado

CPU time: 110 ms indica que el servidor está utilizando el índice no agrupado, lo que resulta en menor uso de CPU ya que los datos no están ordenados físicamente.

Elapsed time: 2635 ms muestra que el tiempo total transcurrido es ligeramente menor que con el índice agrupado, posiblemente debido a que la estructura del índice no agrupado es más sencilla.

Conclusión

Ambos índices ofrecen ventajas y desventajas según el tipo de consulta y el patrón de acceso a los datos:

Índice Agrupado: Es más eficiente para consultas que buscan un rango de valores, ya que los datos están físicamente ordenados. Aunque requiere más CPU, el tiempo total transcurrido es competitivo.

Índice No Agrupado: Ofrece flexibilidad para múltiples consultas específicas y usa menos CPU, pero puede ser menos eficiente para consultas de rango debido a la falta de orden físico en los datos.

En resumen, el índice agrupado puede ser más ventajoso para consultas que requieren acceso a un rango de valores, mientras que el índice no agrupado es útil para consultas específicas en columnas individuales. Los resultados obtenidos están alineados con estas expectativas y muestran cómo cada tipo de índice afecta el rendimiento de las consultas.

\*\*\*\*\*\*\*\*\*\*\*- Eliminar el índice no agrupado en la tabla 'gasto\_sin\_indices'

DROP INDEX IDX\_gasto\_sin\_indices\_fechapago ON gasto\_sin\_indices;\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\*\*\*\*\* Crear índice agrupado en la tabla 'gasto\_sin\_indices'

CREATE CLUSTERED INDEX IDX\_Clustered\_gasto\_sin\_indices\_fechapago ON gasto\_sin\_indices(fechapago;\*\*\*\*\*\*\*\*\*

\*\*\*\*\*\*\*\*\*\*\*\*

---SE VUELVE A COMPARAR LOS TIEMPOS DE ACCESOS AHORA CON EL INDICE AGRUPADO CREADO\*\*\*\*\*\*\*\*\*\*

SET STATISTICS TIME ON;

SET STATISTICS IO ON;

-- Ejecutar la consulta en la tabla original con el índice existente

SELECT \* FROM gasto WHERE fechapago BETWEEN '2023-01-01' AND '2023-02-28';

SET STATISTICS TIME OFF;

SET STATISTICS IO OFF;

SET STATISTICS TIME ON;

Análisis de Resultados

CPU time:

Tabla Original con Índice Existente: 328 ms

Tabla con Índice Agrupado: 109 ms

Elapsed time:

Tabla Original con Índice Existente: 2794 ms

Tabla con Índice Agrupado: 2865 ms

Interpretación

Tabla Original con Índice Existente

CPU time: 328 ms indica un uso más alto de la CPU, posiblemente debido a la optimización que se realiza al usar un índice preexistente.

Elapsed time: 2794 ms sugiere que, aunque la CPU está trabajando más, el tiempo total transcurrido es menor, mostrando una buena eficiencia en la búsqueda.

Tabla con Índice Agrupado

CPU time: 109 ms indica un menor uso de la CPU, lo cual es eficiente para procesos simples o menos carga de trabajo de la CPU.

Elapsed time: 2865 ms muestra que el tiempo total transcurrido es ligeramente mayor, lo cual podría ser porque el índice aún no se ha optimizado por completo.

Conclusión

Comparar estos resultados nos proporciona las siguientes observaciones:

Índice Agrupado:

Ventajas: Menor uso de CPU (109 ms), eficiente para consultas de rango.

Desventajas: Elapsed time (2865 ms) es mayor, posiblemente debido a la optimización pendiente del índice.

Índice No Agrupado Preexistente:

Ventajas: Elapsed time (2794 ms) es menor, indicando eficiencia en la búsqueda con el índice existente.

Desventajas: Mayor uso de CPU (328 ms), lo que puede ser una carga para consultas intensivas en CPU.

En resumen, los índices agrupados pueden ser beneficiosos para ciertas consultas, pero requieren tiempo para optimizarse completamente.