

2022

Optimización de Consultas

LICENCIATURA EN SISTEMAS DE INFORMACIÓN – BASES DE DATOS

Profesores

Pescio Pablo

Schab Esteban

Alumnos

Errandonea Gonzalo

Gimenez Verónica

Grasso Natasha

1. **Para qué sirve el optimizador de consultas.**

Sirve para mejorar los tiempos de respuesta en un sistema de gestión de bases de datos relacional, ya que la optimización es el proceso de modificar un sistema para mejorar su eficiencia o también el uso de los recursos disponibles.

1. **Qué es un plan de ejecución.**

Un plan de ejecución de consulta es una definición de los siguientes elementos: **La secuencia en la que se tiene acceso a las tablas de origen.** **Los métodos que se usan para extraer los datos de cada tabla**. Por lo general, hay métodos diferentes para tener acceso a los datos de cada tabla.

1. **Qué tareas realiza el DBMS para analizar una consulta.**

Pasos para el procesamiento de una consulta:

**Análisis léxico**

* Identificar componentes en el texto de la consulta (SQL)

**Análisis sintáctico**

* Revisar la sintaxis (corrección gramatical).

**Validación semántica**

* Verificar validez de nombres de tablas, vistas, atributos.

**Traducción de la consulta a una representación interna**

* Esto permite que la máquina la manipule mejor, transformar SQL en Álgebra relacional (es lo que generalmente realiza el SGBD en esta etapa).

1. **Qué datos tiene las estadísticas de las tablas de una base de datos.**

Información estadística:

**Para cada tabla**

* Cardinalidad (n° de tuplas).
* Factor de bloqueo (tuplas que caben en un bloque).
* Bloques ocupados.
* Método de acceso primario y otros métodos de acceso (hash, árbol b, etc.)
* Atributos indexados, de dispersión, de ordenamiento

**Para cada atributo**

* Valores de dominios distintos almacenados.
* Valores, máximo, mínimo, etc.

1. **Nombre algunas reglas heurísticas para mejorar una consulta.**

* En la optimización heurística se aplican reglas de transformación para modificar la representación interna de la consulta, a fin de mejor el rendimiento.
* Una consulta en Álgebra Relacional se puede expresar de diferentes maneras, obteniéndose similares resultados
* Así también, un lenguaje de consulta comercial, como SQL, permite realizar la misma consulta de diferentes maneras, pero el rendimiento de la ejecución de la consulta, no debe depender de como sea expresada la misma

1. **Enumere las buenas prácticas para realizar consultas.**

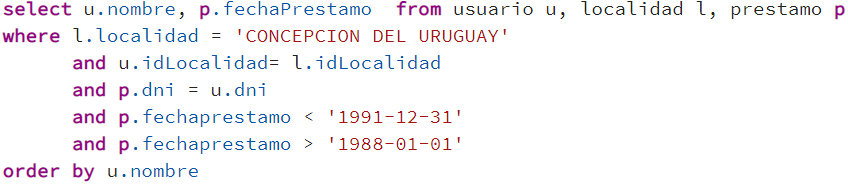
* Ejecutar operaciones de selección σ tan pronto como sea posible.
* Realizar primero las selecciones σ que sean más restrictivas, que generen menor número de columnas.
* Ejecutar las operaciones de proyección Π tan pronto como sea posible.
* Combinar productos cartesianos con una selección, cuya condición representa una condición de reunión, convirtiéndolos en un Join.

1. **Busque en la documentación de postgresql los distintos tipos de índices que se pueden generar y en caso utilizaría cada uno de ellos.**

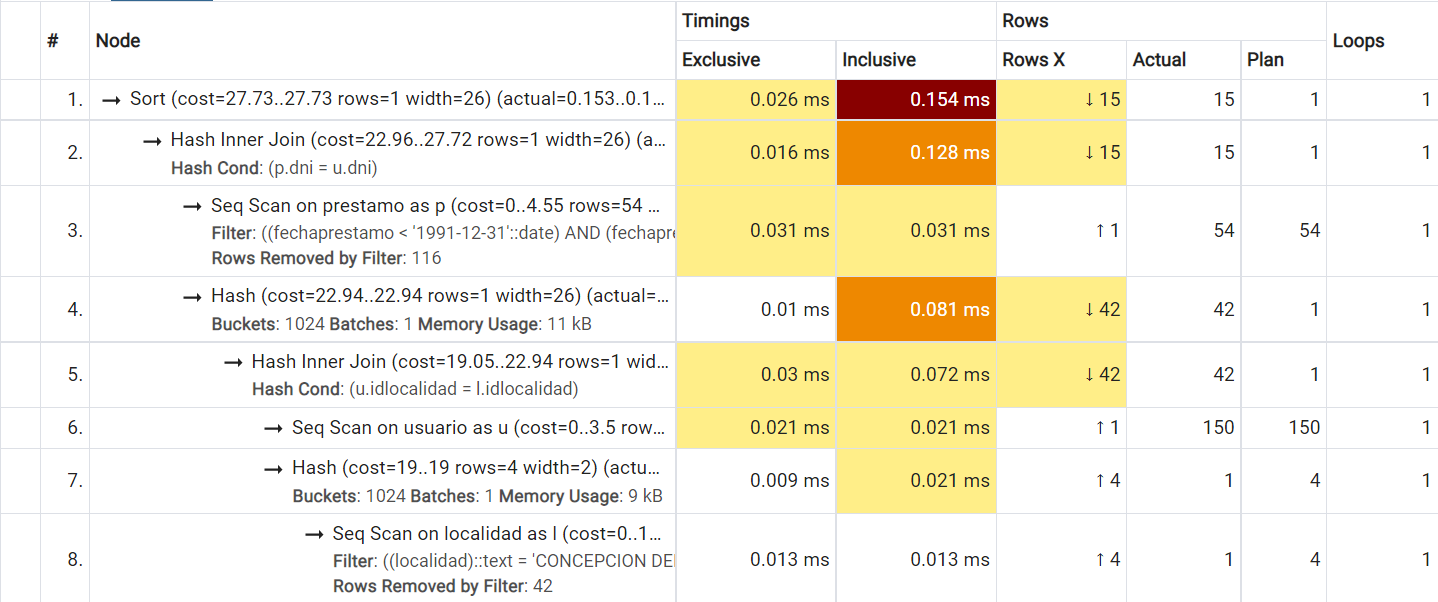
Existen varios tipos de índices: B-tree, Hash, GiST, GIN, BRIN. Cada tipo de índice utiliza un algoritmo diferente que se adapta mejor a diferentes tipos de consultas. De forma predeterminada, el comando CREATE INDEX crea índices de B-tree (Árbol B), que se ajustan a las situaciones más comunes.

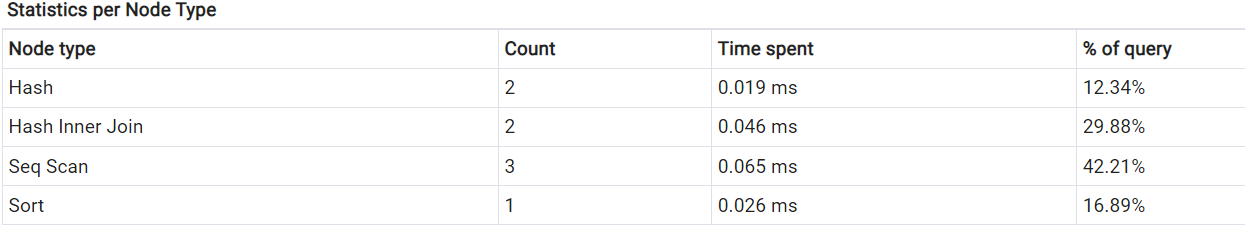
1. **Tomando consultas del ejercicio de elecciones o de libro, el que el grupo prefiera, realice una estimación de costos de consultas más complejas, usando el operador explain en pgadmin para analizar el costo de la consulta, evalúe en grupo si es necesario generar un índice, realícelo y vuelva a ejecutar el explain para ver si mejora la consulta.**

Tomando como ejemplo la siguiente consulta de *libros*:



Obtuvimos los siguientes resultados:

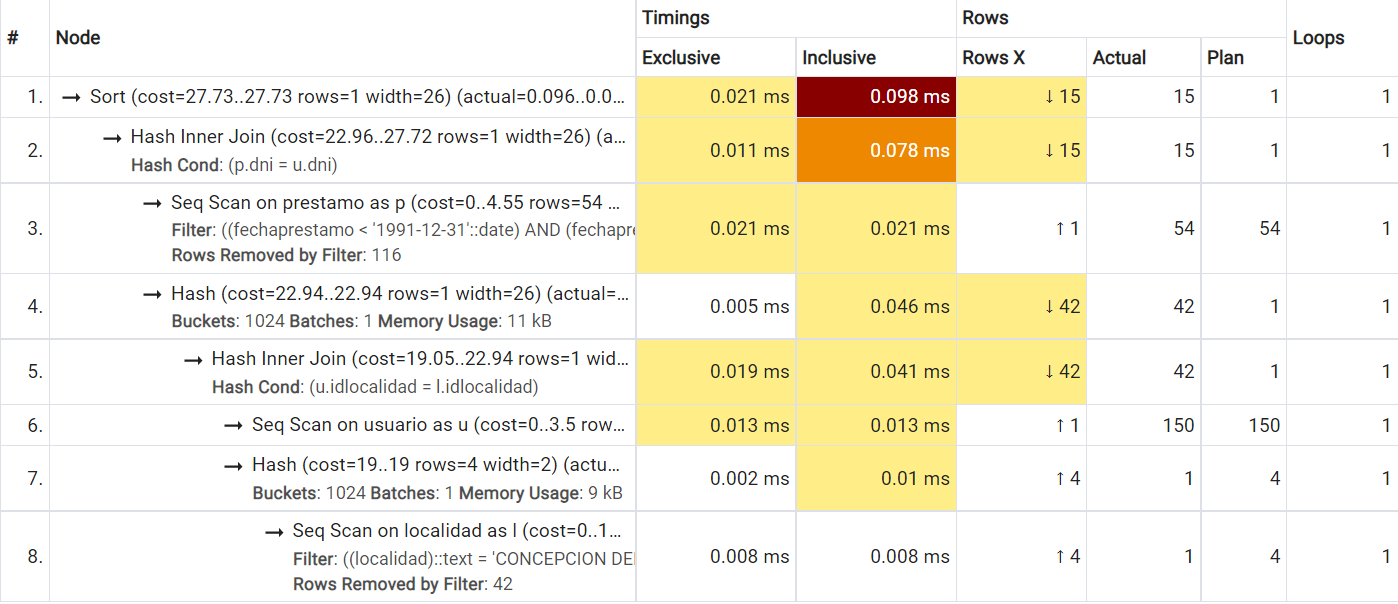


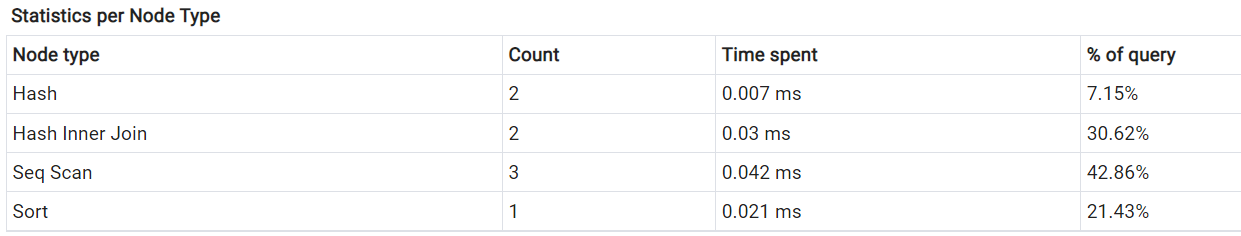


Y luego de crear el siguiente índice:



Obtuvimos los siguientes resultados:





Donde podemos observar que, en efecto, la consulta que utilizó un nuevo índice creado fue más rápida que la anterior.