

Fragmentación horizontal primaria en base de datos relacional, basado en la teoría de Tamer Özsu

La fragmentación se emplea para mejorar el rendimiento, localizando los datos donde más se consultan, los fragmentos son definidos usando condiciones lógicas (predicados) aplicadas a columnas, las condiciones lógicas pueden ser sencillas (mitern, usando sentencias como AND) o complejas y completas (completitud, si se considera que los fragmentos cubran todos los patrones de consulta), puede haber fragmentos vacíos usados para mantener la consistencia y el anexo de datos fuera de rango puede que requieran una nueva fragmentación.

Como diseñar fragmentos en base de datos relacionales / algebra COM_MIN (complete_minimal; todos los posibles fragmentos están cubiertos por los fragmentos / solo se usan los fragmentos relevantes, nada redundante)

El algoritmo COM_MIN genera un conjunto óptimo de condiciones para hacer los fragmentos de la siguiente manera;

- Elige un predicado p_1 que efectivamente divide la tabla de acuerdo con cómo se accede (regla 1)
- Toma una lista de predicados simples, tomando la regla 1 como primer elemento
- Selecciona los predicados relevantes (los que afectan a los accesos) buscando predicados p_j que divida más finamente algún fragmento ya existente en el conjunto y lo agrega si es relevante
- Genera un conjunto final que sea completo y mínimo

Las consideraciones para implementar el algoritmo COM_MIN

1. El acceso por aplicaciones; como las aplicaciones consultan la tabla, por que de lo contrario no se sabe si un predicado es relevante o no
2. Definición de “diferente acceso”; un predicado es relevante si provoca que diferentes aplicaciones accedan a diferentes fragmentos, es decir la frecuencia de acceso por tupla no es igual entre fragmentos lo que justifica la separación
3. Evaluar relevancia de predicados; si un predicado no cambia la forma en que se accede a los datos, no se debe incluir
4. Evitar fragmentos vacíos o contradictorios
5. Complejidad combinatoria; si hay muchos predicados el número de combinaciones posibles (mininterns) crece exponencialmente