

# CC42A - Control 3

Profs. Claudio Gutiérrez y Gonzalo Navarro

21 de Junio de 2002

## Optimización y Planes de Acceso

Sea una tabla  $A$  con 500.000 registros de 128 bytes,  $B$  con 1.000.000 de registros de 256 bytes, y  $C$  con 100.000 registros de 1024 bytes. La página de disco tiene 4096 bytes y tenemos 8 Mb de memoria RAM.  $A$  tiene columnas  $a_1$  (clave, 4 bytes),  $a_2$  (unas 500 distintas, 100 bytes) y  $a_3$  (unas 10.000 distintas, 24 bytes).  $B$  tiene columnas  $b_1$  (40 bytes, unas 100.000 distintas),  $b_2$  (40 bytes, unas 100.000 distintas y forma clave junto con  $b_1$ ), y  $b_3$  (176 bytes, unas 1.000 distintas).  $C$  tiene columnas  $c_1$  a  $c_4$ , todas de 4 bytes, ninguna clave, y unas 1.000 distintas en cada caso; y  $c_5ac_8$ , todas de 252 bytes, ninguna clave, y unas 1.000 distintas en cada caso. Las tablas están almacenadas en disco sin ningún índice ni orden.

Para cada una de las siguientes consultas: (1) conviértala a álgebra relacional, (2) optimízela, (3) proponga planes de acceso alternativos que le parezcan prometedores, considerando incluso la creación de índices de distinto tipo.

1.

```
SELECT a1, b3
FROM A, B, C
WHERE a1 = c1
```

2.

```
SELECT a1, c1, c6
FROM A, B, C
WHERE a1 = c1 AND c1 = c3
```

3.

```
SELECT a1, b3
FROM A, B, C
WHERE a1 = c1
```

4.

```
SELECT a1, b2
FROM A, B, C
WHERE c1 ≤ a1 AND a1 ≤ c4 AND c4 ≤ c3 AND b1 ≠ b2
```

5. Use su imaginación para generar otras.

## Implementación de Operadores Relacionales

Proponga distintas estrategias para ejecutar la operación relacional de unión e intersección, tal como se hizo para los otros operadores relacionales. Analice su costo, y la posibilidad de ordenar o crear índices sólo para realizar esa operación.

## Recuperación

Calcule la frecuencia óptima de checkpointing en función de la frecuencia de las caídas del sistema. Agregue los parámetros que necesite para determinar la frecuencia óptima (ej. cantidad de operaciones de update por segundo, costo de acceder el log, etc.).

Indique ventajas y desventajas de los tres mecanismos de recuperación vistos en clase: logs con updates diferidos, logs con updates inmediatos, y shadow paging.