

## Ejercicio Integrador de Optimización de Consultas

En la terminal de micros de Retiro se cuenta con un sistema con el siguiente esquema de base de datos:

Micro(numMicro, marca, numEmpresa)

Empresa(numEmpresa, nombreEmpresa, fechaCreacion, direccion)

Viaje(numMicro, fechaViaje, destino)

- Micro.numEmpresa no admite nulos y es FK de Empresa.
- Viaje.numMicro es FK de Micro.

Se desea optimizar la siguiente consulta:

```
SELECT
    numMicro, numEmpresa, nombreEmpresa, fechaCreacion, direccion, destino, fechaViaje
FROM
    Viaje V, Empresa E, Micro M
WHERE
    M.numMicro = V.numMicro
AND
    M.numEmpresa = E.numEmpresa
AND
    E.fechaCreacion >= '1/1/2000'
AND
    V.fechaViaje >= '1/1/2007'
AND
    V.fechaViaje < '1/2/2007'
AND
    M.marca = 'Mercedes Benz'
```

Se cuenta con las siguientes estructuras adicionales:

- $I_1$ : Índice Hash sobre *fechaCreacion* en la tabla *Empresa*.
- $I_2$ : Índice B+ unclustered sobre *numEmpresa* en *Empresa* (altura 3).
- $I_3$ : Índice B+ clustered sobre *<fechaViaje, numMicro>* en *Viaje* (altura 5).
- Las tablas *Micro* y *Empresa* se guardan como heap file.

La longitud de los campos es de 128 bytes cada uno.

Se tiene como datos:

- Tuplas Micro: 200.000.
- Tuplas Empresa: 3.000.
- Tuplas Viaje: 2.500.000.
- Micros con marca 'Mercedes Benz': 50.000.
- Viajes de enero del 2007: 100.000.
- Empresas creadas desde el año 2000: 2.000.

Además:

- Se cuenta con 5 bloques de memoria.
- El tamaño de bloque es 2048 bytes.
- Se asume distribución uniforme.

- a) Armar el árbol canónico de la consulta.
- b) Obtener un árbol optimizado, indicando su plan de ejecución, y calcular los costos.
- c) Describir brevemente qué estructura agregaría para mejorar la performance de la consulta (sin calcular costos).