

INDICES

Considere una relación E con atributos id, nombre, edad y salario. La tabla E contiene 1000 tuplas y id es la clave primaria de la relación. Asumimos que el menor id es 1 y los demás ids están en orden secuencial creciente (sin saltar ningún valor). Todos los registros en la tabla E están almacenados secuencialmente según el id (clustering index). En el sistema, los *bloques de datos* almacenan hasta 10 registros o tuplas (asumimos que las tuplas son de tamaño fijo), y los *bloques de índices* almacenan hasta 20 entradas (valor, puntero) y las entradas del índice están ordenados por valor. Para simplificar consideraremos que el sistema utiliza una búsqueda secuencial sobre el índice, comenzando desde el primer bloque de índice por cada consulta.

id	nombre	edad	salario
1	N1	21	\$8100
2	N2	22	\$4532
...			
10	N10	19	\$6800

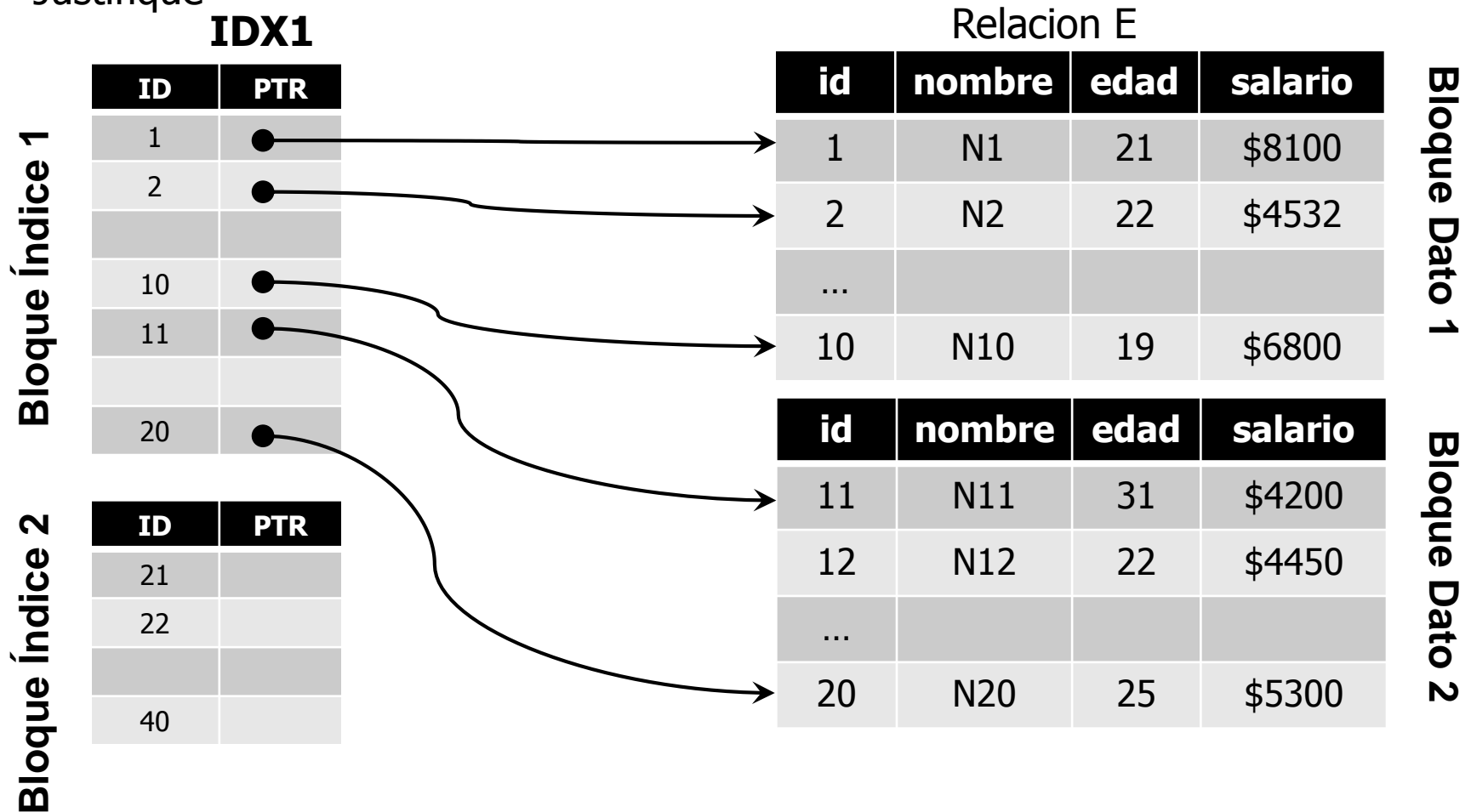
Bloque de datos 1

id	nombre	edad	salario
11	N11	31	\$4200
12	N12	22	\$4450
...			
20	N20	25	\$5300

Bloque de datos 2

INDICES

Suponga que existe un índice **denso** IDX1 sobre E.id. ¿Cuántos bloques (cantidad de IOs) debe leer el sistema para obtener el contenido del registro con **id=530**? Justifique



INDICES

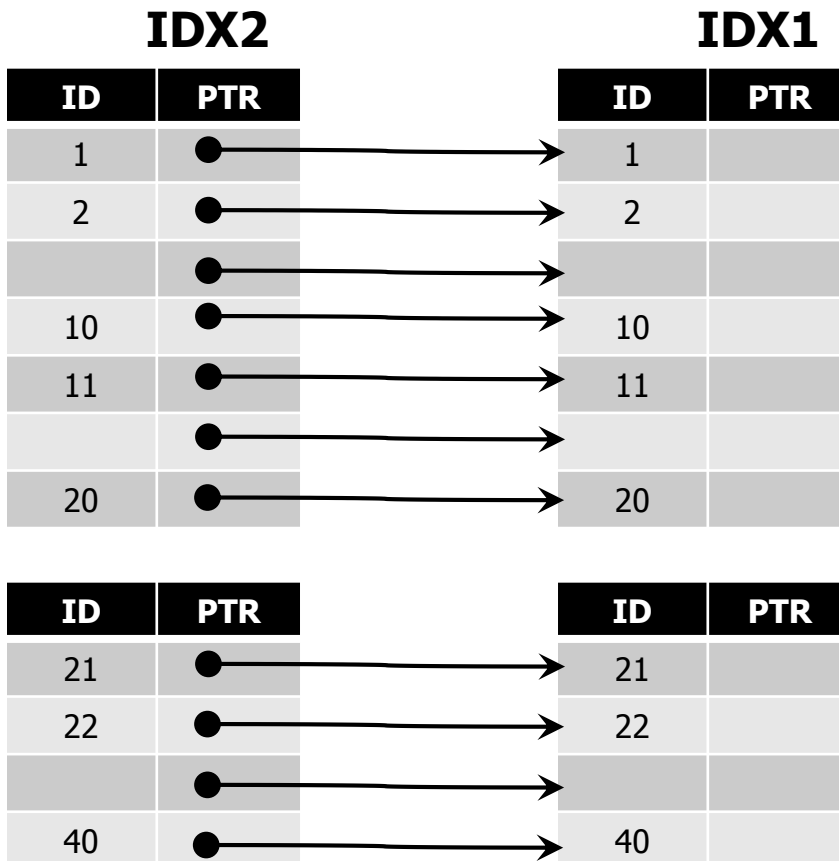
Suponga que existe un índice **denso** IDX1 sobre E.id. ¿Cuántos bloques (cantidad de IOs) debe leer el sistema para obtener el contenido del registro con **id=530**? Justifique

$$\left\lceil \frac{530}{20} \right\rceil = 27 \quad \text{Bloques de índices}$$
$$+ 1 \quad \text{Bloques de datos}$$
$$\left. \begin{array}{l} 27 \\ + 1 \end{array} \right\} 28 \quad \text{Bloques}$$

INDICES

Se desea mejorar la velocidad creando un índice IDX2 sobre IDX1. ¿Qué tipo de índice debería ser IDX2, ralo (sparse) o denso? Justifique su respuesta.

¿Cuánto ocuparía IDX2 si fuese denso?



Ocuparía lo mismo que IDX1.

Agregaría un nivel de indirección extra debiendo traer un bloque más.

Tardaría más que IDX1 solo.

INDICES

¿Cuántos bloques (cantidad de IOs) se deben leer para obtener el contenido del registro con id=530 si se utiliza el IDX2? Justifique

2 bloques de IDX2
+ 1 bloque de IDX1
+ 1 bloque de datos

4 bloques

$$\left\lceil \left\lceil \frac{530}{20} \right\rceil + 1 \right\rceil + 1 + 1 = 4$$

Id=530

IDX2

ID	PTR
1	
21	
381	
401	
521	
781	
801	
981	
null	

IDX1

ID	PTR
1	
2	
10	
11	
20	
21	
22	
40	



INDICES

Se desea construir otro índice pero sobre el atributo **E.salario**. ¿Qué tipo de índice se debe crear? ¿Primario o **secundario**? ¿Ralo o **denso**? Justifique

Suponga que existen 50 valores distintos de salarios y que existen una igual cantidad de registros para cada valor de salario.

¿Qué cantidad de registros existen por cada salario?

$$1000 / 50 = 20 \text{ registros (un bloque de índice completo)}$$

¿Cuántos bloques tiene el índice? $1000 / 20 \text{ (tuplas)} = 50 \text{ bloques}$

Asuma que los valores repetidos se almacenan en el índice y el índice se encuentra ordenado por salario.

¿En que posición del índice se encuentran los registros correspondientes al mayor salario?

Al final. En el último bloque.

INDICES

IDX3

salario	PTR
4200	
4200	

Bloque Índice 1
(20 valores)

salario	PTR
4450	
4450	

salario	PTR
9700	
9700	

Bloque Índice 50
(20 valores)

Relacion E

id	nombre	edad	salario
1	N1	21	\$8100
2	N2	22	\$4532
...			
10	N10	19	\$6800

id	nombre	edad	salario
11	N11	31	\$4200
12	N12	22	\$4450
...			
20	N20	25	\$5300

Bloque Dato 1

Bloque Dato 2

¿Cuántos bloques deben ser leídos para obtener el contenido de todas las tuplas que tengan el mayor salario? Justifique

50 (bl. índices) + 20 bl. datos = 70 bloques

INDICES

Si suponemos que el índice del inciso e no contiene valores repetidos sino que cada valor apunta a un conjunto de bloques contiguos que contienen los punteros a los bloques de datos. ¿Cuántos bloques deben ser leídos para obtener el contenido de todas las tuplas que tengan el mayor salario? Justifique

