Observaciones lab 6

Juan Andrés Eslava Tovar 202012035

David Alejandro Fuquen Flórez 202021113

a) Teniendo en cuenta cada uno de los requerimientos ¿Cuántos índices implementaría en el Reto? y ¿Por qué?

Para saber cuántos índices implementaría en el reto, veremos cuantos son necesarios para cada requerimiento.

Req1: Índice por año de nacimiento de artista.

Req2: Índice por año de adquisición de obra.

Req3: Índice por ConstituentID de las obras (es decir, por artista). Índice por nombre del artrista (llave: ConstituentID, valor: nombre del artista).

Req4: Índice por nacionalidad de la obra. Índice de nacionalidad por artista (llave: ConstituentID, valor: Nacionalidad del artista)

Req5: Índice de obras por departamento del museo.

b) Según los índices propuestos ¿en qué caso usaría Linear Probing o Separate Chaining en estos índices? y ¿Por qué?

Para cada índice, nos basaremos en la complejidad temporal teórica explicada en clase. En ese orden de ideas, usaremos Separate Chaining en cada caso, pues, aunque en casi todos los casos tiene la misma complejidad temporal, en el peor caso sin rehash tiene una complejidad de O(N/M), mientras que Linear Probing tiene una mayor complejidad temporal de O(N).

c) Dado el número de elementos de los archivos MoMA, ¿Cuál sería el factor de carga para estos índices según su mecanismo de colisión?

Como es un separate chaining, se eligirá un factor de carga de 4.0. Es decir, que se aceptan por promedio 4 tuplas llave valor por cada casilla de la tabla Hash.

Paso 3 y 4:

# **Ambientes de pruebas**

|  |  |  |
| --- | --- | --- |
|  | **Máquina 1** | **Máquina 2** |
| **Procesadores** | Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz 2.11 GHz | AMD Ryzen 5 3500U with Radeon Vega Mobile Gfx 2.10 GHz |
| **Memoria RAM (GB)** | 8.00 GB | 8,00 GB |
| **Sistema Operativo** | Windows 10 Pro 64-bits | Windows 10 Home Single Language 64-bits |

*Tabla 1. Especificaciones de las máquinas para ejecutar las pruebas de rendimiento.*

*Máquina 1:*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Small | 5pct | 20pct | 80pct |
| SC 4 | 62,5 | 357 | 1250 | 4406,25 |
| SC 2 | 62,5 | 343,75 | 1343,75 | 4750 |
| SC 8 | 62,5 | 421,7 | 1062,5 | 5062,5 |
| LP 0.5 | 46,88 | 406,25 | 1093,75 | 4265,63 |
| LP 0.2 | 62,5 | 359,38 | 1265,625 | 4328,13 |
| LP 0.8 | 62,5 | 390,75 | 1125 | 4609,34 |

*Máquina 2:*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Small | 5pct | 20pct | 80pct |
| SC 4 | 62,5 | 468,75 | 1765,625 | 6671,875 |
| SC 2 | 62,5 | 484,375 | 1750 | 6.471,875 |
| SC 8 | 78,125 | 500 | 1906,25 | 7250 |
| LP 0.5 | 78,125 | 453,125 | 1593,75 | 6046,875 |
| LP 0.2 | 78,125 | 453,125 | 1531,25 | 6437,5 |
| LP 0.8 | 78,125 | 578,125 | 1718,75 | 6578,125 |

C) Al evaluar evaluar los archivos que aparecen en las tablas en ambas máquinas, el Linear Probing con un factor de carga de 0.50 tiende a ser más rápido que el Separate Chaining con factor de carga de 4.00.

D) Tras analizar los resultados, decidimos utilizar Separate Chaining, con factor de carga de 4.0, con 10000 elementos iniciales para el índice de técnicas o medios. Pues, aunque existen Linear Probings con rapidez de carga un poco más rápida, Separate Chaining es más eficiente en el uso de recursos. Por ende, no vale la pena usar Linear Probing por tan poca diferencia de tiempo de carga.

E) Tras analizar los resultados, decidimos utilizar Separate Chaining, con factor de carga de 4.0, con 10000 elementos iniciales para el índice de nacionalidades. Pues, aunque existen Linear Probings con rapidez de carga un poco más rápida, Separate Chaining es más eficiente en el uso de recursos. Por ende, no vale la pena usar Linear Probing por tan poca diferencia de tiempo de carga.