DOCUMENTO ANÁLISIS

“NOTA IMPORTANTE: Debido a la crisis que está viviendo el país y ya que no poseo un computador con la capacidad de cargar el archivo grande con los 527000 datos, tuve que realizar la toma de datos con el archivo que tiene solo 50000 datos el cual se encuentra en:

Proyecto 2/Comparendos\_DEI\_2018\_Bogotá\_D.C\_50000\_.geojson” Gracias por su atención.

Carga de datos:

Después de realizar la carga de los datos podemos capturar la siguiente información

|  |  |  |
| --- | --- | --- |
|  | Tabla de Hash Linear  Probing | Tabla de Hash Separate Chaining |
| Número de duplas (K, V) en la tabla (valor N) | 16181 | 16181 |
| Tamaño inicial del arreglo de la tabla (valor M inicial) | 5 | 5 |
| Tamaño final del arreglo de la tabla (valor M final) | 25717 | 6421 |
| Factor de carga final (N/M) | 0,63 | 02,52 |
| Número de rehashes que tuvo la tabla (desde que se creó) | 12 | 10 |

Con estos datos podemos afirmar que el crecimiento de la Tabla de Hash Separate Chaining es mucho menor que la Tabla de Hash Linear Probing pues mientras la primera tiene la necesidad de hacer solo 10 rehashes la segunda tiene que hacer 12. Esto hace que los tamaños finales de los arreglos de la tabla tengan tamaños bastantes diferentes pues en una Tabla de Hash Separate Chaining el tamaño final es de 6421 mientras que en una Tabla de Hash Linear Probing el tamaño final es de 25717.

Pruebas de desempeño:

Después de hacer las pruebas de desempeño varias veces conseguimos la siguiente información

|  |  |  |
| --- | --- | --- |
|  | Tabla de Hash Linear  Probing | Tabla de Hash Separate Chaining |
| Tiempo mínimo de get(…) | 7 | 5 |
| Tiempo promedio de get(…) | 10 | 8 |
| Tiempo máximo de get(...) | 14 | 11 |

Después de esta información podemos ver que en una Tabla de Hash Separate Chaining se busca mucho mejor una llave determinada.