

TDE 03 - Ordenação RA 04

Renan Antonio Hammerschmidt Krefta¹

¹Pontifícia Universidade Católica do Paraná (PUCPR)

renankrefta@hotmail.com

Abstract. *This document presents comparison and analysis between three simple Hash Functions and presents line graphics and tables containing the comparison data.*

Resumo. *Este documento apresenta comparação e análise entre três funções de hash simples e apresenta gráficos e tabelas contendo os dados de comparação*

1. Descrição do Código

Foram implementados no código três funções simples de hash, a primeira função calcula o hash com base no resto da divisão pelo tamanho da tabela, a segunda função calcula com base na divisão do dobro do elemento pelo tamanho da tabela e por fim, a terceira função realiza seu cálculo multiplicando a chave por uma constante qualquer (0.41215), e então multiplica o resultado da subtração desse produto pelo casting (int) desse mesmo produto, por exemplo, digamos que um produto seja 1.3, seu casting será 1 e a subtração deste produto pelo seu casting retornará $1.3 - 1 = 0.3$, multiplicando isso pelo tamanho da tabela (100, por exemplo) retornará como índice $100 * 0.3 = 30$. Foi utilizado Rehashing para tratar colisões no código e quando uma tabela está cheia, ela não receberá mais inserções.

As funções hash foram comparadas das seguintes formas: Quantidade de colisões; Tempo levado para as inserções; Quantidade de comparações e tempo levado para as buscas.

Todos os dados são resultados de uma média dada por 5 execuções de cada função e os tempos são medidos em nanosegundos (ns).

Link Para o Código: <https://github.com/Ziminny1/Avaliacao-RA03-Estruturas-De-Dados>

2. Tabelas

Alguns dos campos de tempo de busca e de comparações das tabelas estão com N/A (Não Avaliado) por conta do quão demorado o código estava ficando (mais que 40 minutos), portanto, precisei deixar alguns campos desta forma para que pudesse dar continuidade aos testes.

2.1. Colisoes

Table 1. Colisoes - 20 mil

Tamanho	Resto	Dobra	Multi
10	9	13	25
100	590	630	506
1000	17188	28876	14811
10000	310635	436270	338478
100000	0	0	0

Table 2. Colisoes - 100 mil

Tamanho	Resto	Dobra	Multi
10	18	21	19
100	442	434	444
1000	19884	16268	20997
10000	575368	455736	505512
100000	0	50000	228000

Table 3. Colisoes - 500 mil

Tamanho	Resto	Dobra	Multi
10	11	7	18
100	659	768	504
1000	15525	17550	17683
10000	517996	520992	421825
100000	17454687	11759374	25452760

Table 4. Colisoes - 1 milhao

Tamanho	Resto	Dobra	Multi
10	9	13	3
100	659	568	331
1000	15645	8790	17559
10000	402178	379356	497936
100000	10454210	13358420	17793200

Table 5. Colisoes - 5 milhoes

Tamanho	Resto	Dobra	Multi
10	13	11	5
100	525	600	455
1000	14567	8634	13529
10000	467165	579330	391823
100000	28053253	17556506	15949973

2.2. Tempo de Inserção

Table 6. Tempo de Inserção - 20 mil

Tamanho	Resto	Dobra	Multi
10	394600	613060	582359
100	293379	331680	327800
1000	583340	639420	614119
10000	2886740	4662479	3033299
100000	681559	624299	723900

Table 7. Tempo de Inserção - 100 mil

Tamanho	Resto	Dobra	Multi
10	894140	1277299	1606160
100	1025320	1193179	1407139
1000	1545620	1483419	1974040
10000	5746460	5135939	4928080
100000	3806060	5057819	5539360

Table 8. Tempo de Inserção - 500 mil

Tamanho	Resto	Dobra	Multi
10	4713840	6009320	6355920
100	4860060	5582959	6670460
1000	5686559	5911840	7440700
10000	8941980	9777920	10252280
100000	137493760	91840459	190561260

Table 9. Tempo de Inserção - 1 milhão

Tamanho	Resto	Dobra	Multi
10	9918540	8961359	9047359
100	10455800	9461320	9726839
1000	11133580	11075139	11480120
10000	13398820	13684520	14454780
100000	85457519	102640560	135386619

Table 10. Tempo de Inserção - 5 milhões

Tamanho	Resto	Dobra	Multi
10	44953679	46316060	45951740
100	41767719	43338460	43870380
1000	55068560	50877099	51875099
10000	58509860	48856980	46763700
100000	255682819	176813180	166059400

2.3. Comparações

Table 11. Comparacoes - 20 mil

Tamanho	Resto	Dobra	Multi
10	219899	219903	219915
100	2010490	2010530	2010406
1000	19036188	19047876	19033811
10000	100320635	100446270	100348478
100000	0	0	0

Table 12. Comparacoes - 100 mil

Tamanho	Resto	Dobra	Multi
10	1099908	1099911	1099909
100	10090342	10090334	10090344
1000	99118884	99115268	99119997
10000	900665368	900545736	900595512
100000	0	50000	228000

Table 13. Comparacoes - 500 mil

Tamanho	Resto	Dobra	Multi
10	5499901	5499897	5499908
100	50490559	50490668	50490404
1000	499514525	499516550	499516683
10000	606040700	606043696	605944529
100000	N/A	N/A	N/A

Table 14. Comparacoes - 1 milhão

Tamanho	Resto	Dobra	Multi
10	10999899	10999903	10999893
100	100990559	100990468	100990231
1000	1000014645	1000007790	1000016559
10000	1311457586	1311434764	1311553344
100000	N/A	N/A	N/A

Table 15. Comparacoes - 5 milhões

Tamanho	Resto	Dobra	Multi
10	54999903	54999901	54999895
100	504990425	504990500	504990355
1000	709046271	709040338	709045233
10000	N/A	N/A	N/A
100000	N/A	N/A	N/A

2.4. Tempo de Busca

Table 16. Tempo de Busca - 20 mil

Tamanho	Resto	Dobra	Multi
10	3014600	6264200	7048100
100	15842300	16702900	15908500
1000	138814800	139921900	138458200
10000	721828500	731952000	720975300
100000	1134000	2235400	2678900

Table 17. Tempo de Busca - 100 mil

Tamanho	Resto	Dobra	Multi
10	9818300	9461600	8790600
100	76938400	76696300	76069600
1000	735904900	723358100	727031000
10000	6459583000	6499560800	6475033800
100000	2981900	4437800	7732700

Table 18. Tempo de Busca - 500 mil

Tamanho	Resto	Dobra	Multi
10	46387600	46244800	48163200
100	374381100	374784500	376628800
1000	3556214600	3597674400	3607365900
10000	34929678700	35062668900	35015741400
100000	N/A	N/A	N/A

Table 19. Tempo de Busca - 1 milhão

Tamanho	Resto	Dobra	Multi
10	102654080	99696500	99751140
100	765856700	767433600	772836200
1000	7198374880	7195721500	7235168020
10000	70280022700	71172074700	71071085100
100000	N/A	N/A	N/A

Table 20. Tempo de Busca - 5 milhões

Tamanho	Resto	Dobra	Multi
10	563243620	560887240	569994460
100	3857866500	3863894260	3908466860
1000	36338925500	36365625900	36231232700
10000	N/A	N/A	N/A
100000	N/A	N/A	N/A

3. Graficos

3.1. Colisões

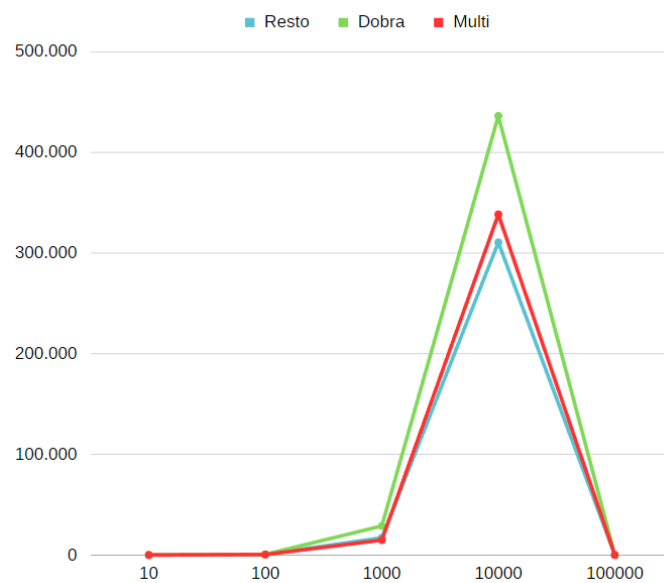


Figure 1. Colisões - 20 mil

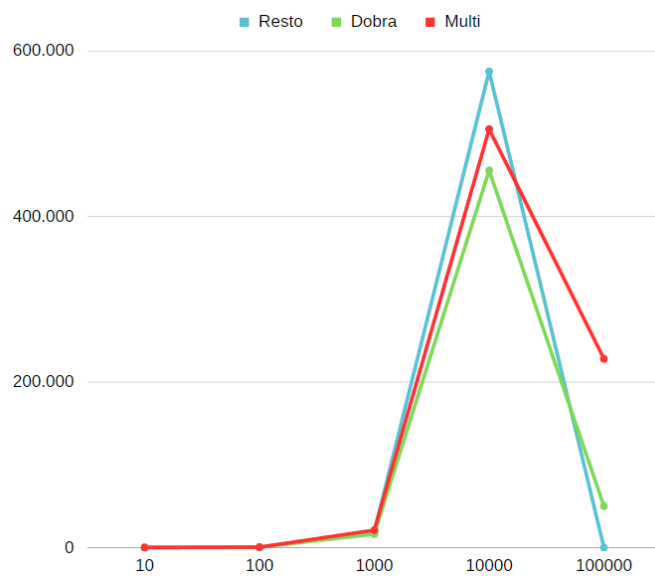


Figure 2. Colisões - 100 mil

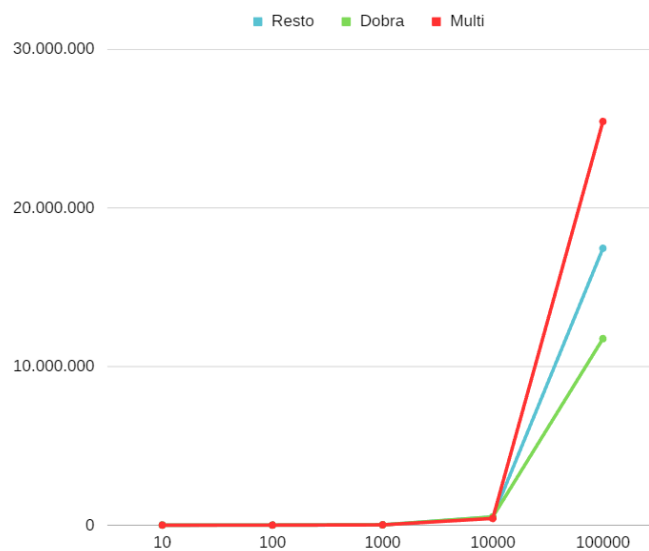


Figure 3. Colisões - 500 mil

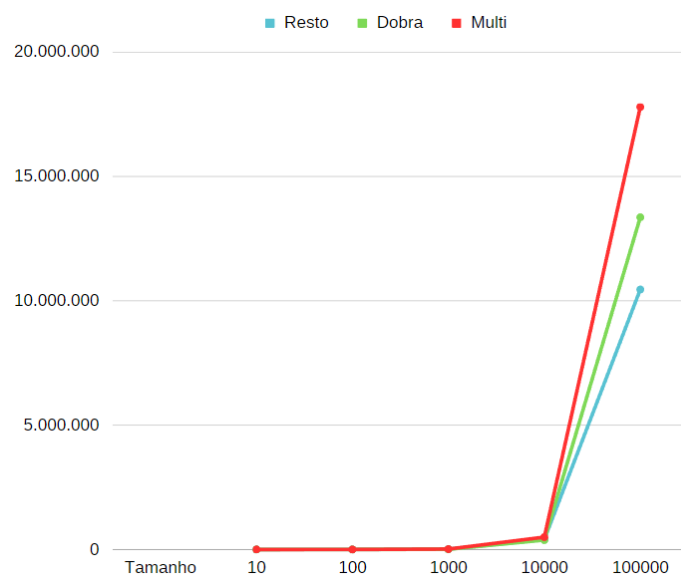


Figure 4. Colisões - 1 milhão

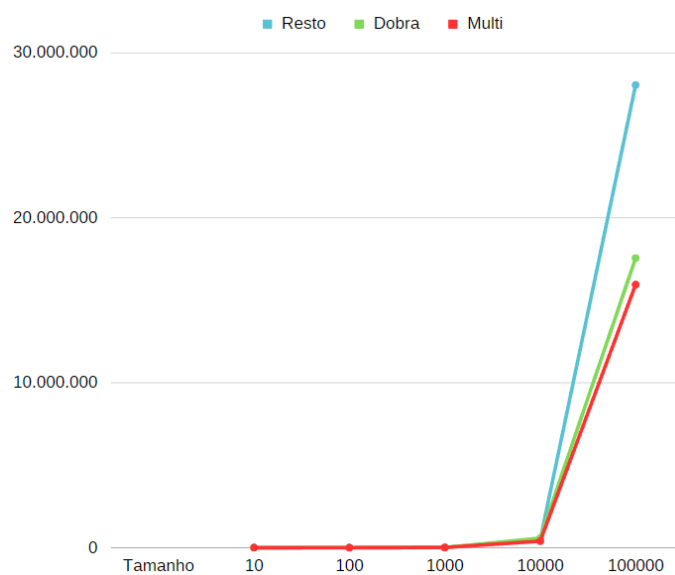


Figure 5. Colisões - 5 milhões

3.2. Tempo de Inserção

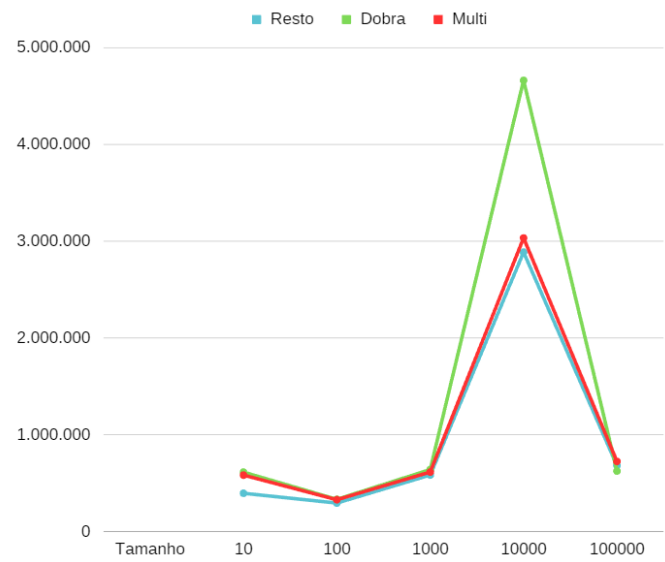


Figure 6. Tempo de Inserção - 20 mil

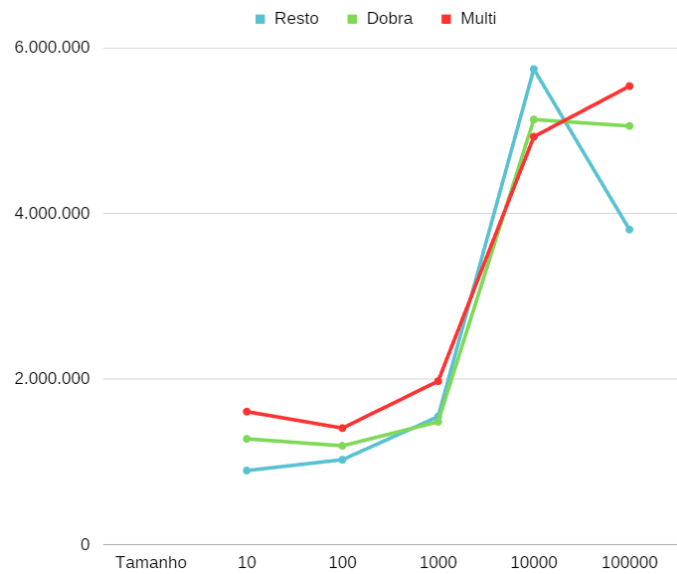


Figure 7. Tempo de Inserção - 100 mil

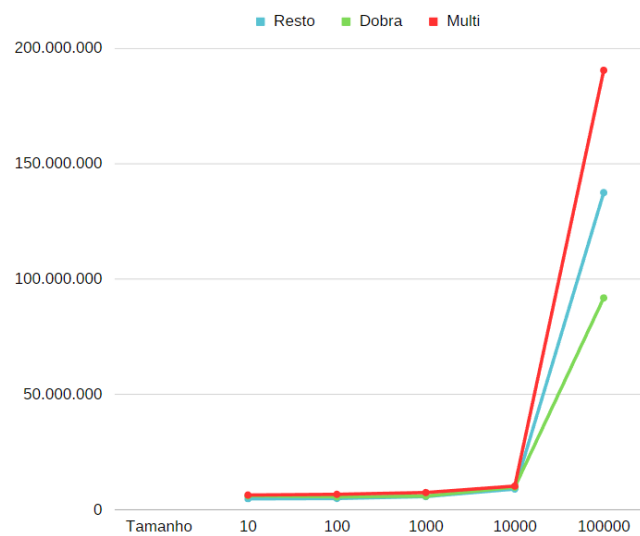


Figure 8. Tempo de Inserção - 500 mil

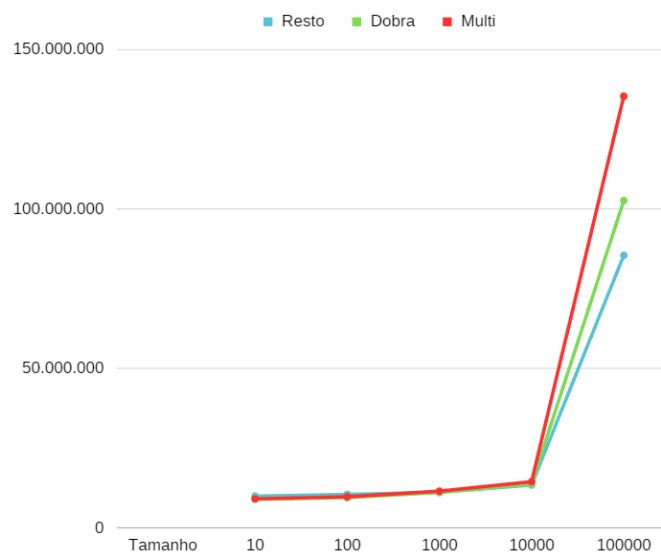


Figure 9. Tempo de Inserção - 1 milhão

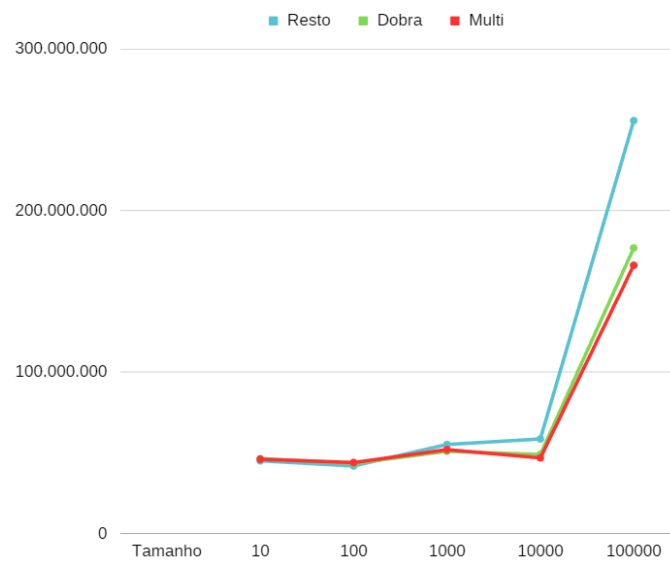


Figure 10. Tempo de Inserção - 5 milhões

3.3. Comparações

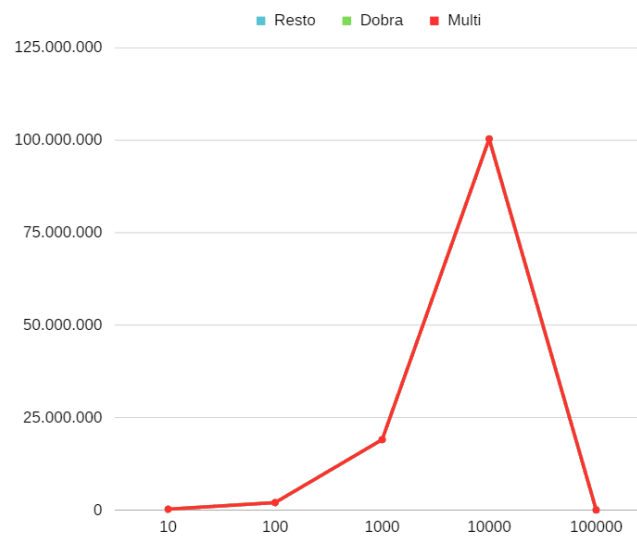


Figure 11. Comparações - 20 mil

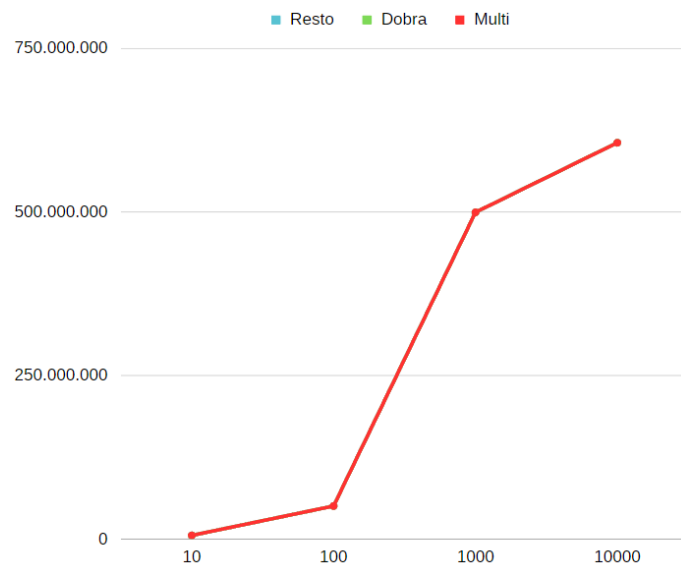


Figure 12. Comparações - 500 mil

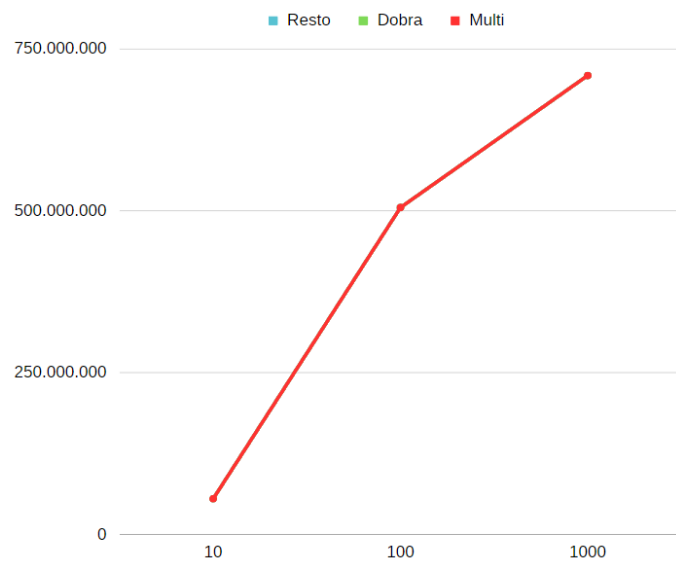


Figure 13. Comparações - 5 milhões

3.4. Tempo de Busca

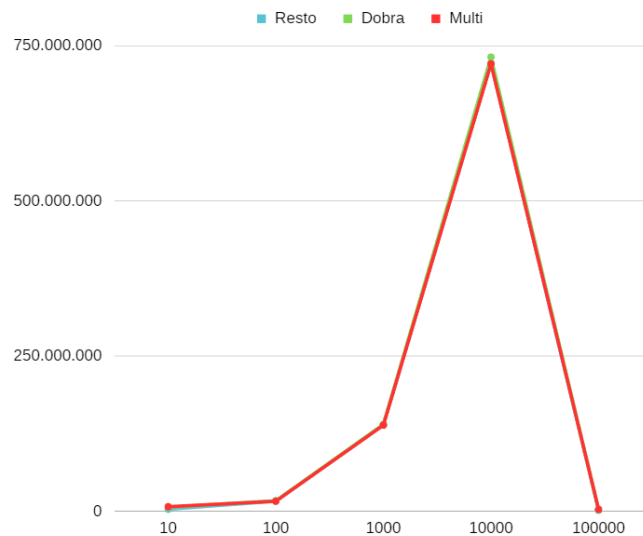


Figure 14. Tempo de Busca - 20 mil

4. Conclusão

Os dados extraídos dos testes revelaram que o tamanho das tabelas em relação ao tamanho do vetor de elementos tem um papel significativo no desempenho quanto ao tempo, número de colisões e comparações, como pode ser visto, principalmente, nas tabelas (Table. 1 Colisoes - 20 mil), (Table. 2 Colisões - 100 mil), (Table. 6 Tempo de Inserção - 20 mil) e (Table. 7 Tempo de Inserção - 100 mil), onde é evidente que, quando o tamanho da tabela é maior ou igual ao tamanho do conjunto de dados em questão, há uma grande diferença de tempo e colisões porque cada elemento sendo inserido terá pouca chance de acabar em um índice já ocupado na tabela, como pode ser visto no caso da função de Resto, onde não demonstra nenhuma colisão e nenhuma comparação necessária, isso também afetou o tempo total que essa função levou para encerrar os testes quando em comparação com as outras duas. Contudo, no que diz respeito às comparações das funções escolhidas, os dados são muito próximos e oscilantes para chegar em um resultado definitivo, mas isso é esperado visto que eram funções simples.