Assignment 6 Write-Up

James Contini

2022 November 20

Description

In this write-up I am discuss the results of this assignment.

1 Investigating Bloom filter bits examined per miss

In this section we will look at how changing the size of the Bloom Filter affects the bits examined per miss. In this figure below is a plot of the bloom filter against the "bepm" *bits examined per miss*. In this

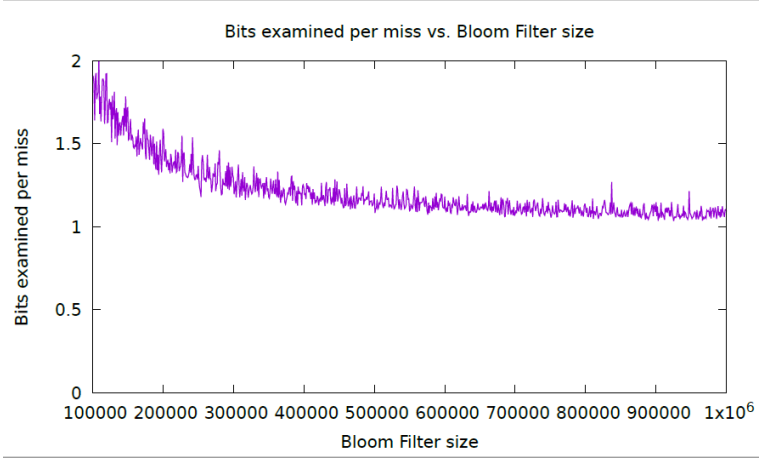

Figure 1: Bits examined per miss plotted against bloom filter size

plot the size of the bloom filter increases by one thousand for every point on the line. The first thing I noticed when making my first observations was the fact that when the bloom filter had a full load the *bepm* evaluated to zero. I think this suggests, along with the shape of the graph that there is a vertical asymptote on the left side going up, which means for values of bloom filter size that where the load approaches one, the left increases forever.

2 Investigating how the Bloom filter size affects the number of lookups performed in the hash table

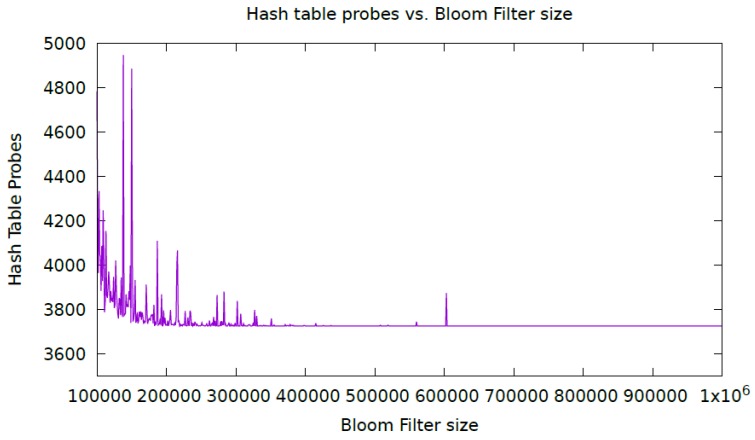

Figure 2: Increments by one thousand per line point *(hashtable lookups vs bloom filter size)*

As the bloom filter size increases the probes of the hash table find an asymptote, and this is only possible when the input is the same. In a perfect world, there should on be as many hash table probes as there are elements in the hash table because that would mean the bloom filter could say for sure every time when the table should be accessed. But since bloom filters are a probabilistic data structure there may be more probes for when the bloom filter is less likely to give a correct answer.

Many values on the left before the asymptote probably resulted in a miss because being on the asymtote means that there were as many probes as were possibly needed.