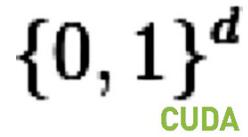


Very fuzzy searching with CUDA

Description	Searching without datastructures.
Period	Fall 2015
Languages & Libs	C++, CUDA
Tags	Thrust, Databases, GitHub, Stack Overflow
GitHub	nikonyrh/stackoverflow-scripts



This is an alternative answer to the question I encountered at [Stack Overflow](#) about fuzzy searching of hashes on Elasticsearch. My original answer used [locality-sensitive hashing](#). Superior speed and simple implementation were gained by using nVidia's [CUDA](#) via [Thrust library](#).

Actually this project didn't need more than a few dozen lines of code to solve the problem, most code is about benchmarking the code with varying problem sizes. Hamming distance was calculated as the Hamming weight of the two 64-bit integers, the used algorithm is "popcount_3" of [Wikipedia article](#). Interestingly it was more efficient to calculate the sum of two 32-bit "halves" weights instead of calculating it for all 64-bits at once, it would take more benchmarking and profiling to find out why this was the case.

Anyway, my laptop's GeForce GT 650M was able to scan through 700 million 64-bit hashes / second. This was far superior performance that I got from Elasticsearch which would need 100 milliseconds to search through 1 million documents but for CUDA it takes only 2.5 milliseconds! The performance difference gets even greater when more documents are searched, for example for 7 million images it took 6 - 7 seconds for ES but only 10 milliseconds for CUDA. Also with this solution we get 100% accuracy on results.

Kernel launch incurs some overhead but it diminishes when searching on more than 5 million images.

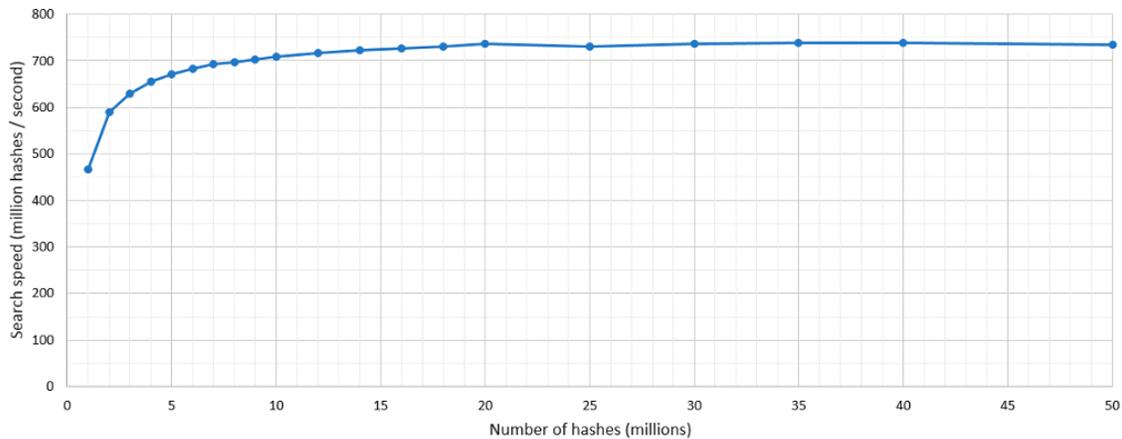


Figure 1: Search performance of 700 million images / seconds was reached.

The CUDA solution could be made HTTP-accessible by for example implementing a FastCGI interface to the script and having a Nginx web server routing requests there. This is very easy to set-up especially if image hashes don't change often and fit in single machine's GPU's memory.