GateKeeper-GPU: Fast and Accurate Pre-Alignment Filtering in Short Read Mapping

Zülal Bingöl	Mohammed Alser	Onur Mutlu	Ozcan Ozturk	Can Alkan
Bilkent University	*ETH Zürich*	*ETH Zürich*	*Bilkent University*	*Bilkent University*
zulal.bingol@bilkent.edu.tr	alserm@inf.ethz.ch	onur.mutlu@inf.ethz.ch	ozturk@cs.bilkent.edu.tr	calkan@cs.bilkent.edu.tr

Abstract—We introduce GateKeeper-GPU, a fast and accurate pre-alignment filter that efficiently reduces the need for expensive sequence alignment. GateKeeper-GPU improves the filtering accuracy of GateKeeper, and by exploiting the massive parallelism provided by GPU threads it concurrently examines numerous sequence pairs rapidly. GateKeeper-GPU is available at https://github.com/BilkentCompGen/GateKeeper-GPU. Please refer to the preprint at *arXiv:2103.14978* for more information.

Index Terms—read mapping, pre-alignment filtering, GPGPU, sequence alignment acceleration

I. INTRODUCTION

The compute-intensive nature of dynamic programming solutions in the verification step of read mapping creates a bottleneck. Our goal is to significantly decrease the execution time of sequence alignment by reducing the workload so that the pipeline visits verification as rarely as possible, and eliminating the candidate locations that exceed a predefined error threshold with a fast and accurate pre-alignment filtering.

II. METHODS

GateKeeper-GPU starts by collecting information on system specifications beforehand to allocate memory wisely and enable the extra features accordingly. Then, data buffers are allocated in unified memory for an optimized data transfer between host and device. Reads and candidate reference segments are encoded in 2-bits for the bitwise algorithm, and finally, filtration is applied on each sequence pair by GateKeeper-GPU kernel with an improved GateKeeper algorithm. GateKeeper-GPU processes the reads exhibiting data-parallel fashion. Since filtering one pair is a simple operation, introducing intra-sequence parallelism would introduce extra overhead due to dependencies. Therefore, each GPU thread completes a single GateKeeper operation. We also integrate GateKeeper-GPU into mrFAST [1] in order to observe practical performance gains.

III. RESULTS

GateKeeper-GPU can filter up to $3\times$ and $20\times$ more read and reference segment pairs than its CPU counterpart, with single and 8 NVIDIA GeForce GTX 1080 Ti GPUs, respectively. When GateKeeper-GPU is combined with a full mapper (mr-FAST) for 100bp reads, it can accelerate the verification stage up to $2.9\times$, and this reflects on the overall performance as up to $1.4\times$ speedup.

Up to $\sim 3\%$ error threshold of the read length, GateKeeper-GPU can correctly reject more than 90% of the potential mappings seeded by mrFAST, with less than 10% of false accept ratio and with no false reject. Its improved algorithm can produce up to $52\times$ less number of false accepts when compared to original GateKeeper [2]. Furthermore, it can filter out all dissimilar sequence pairs produced by Minimap2 with no false accepts for exact matching, and can present a true reject rate of up to 98% on BWA-MEM's candidate mappings.
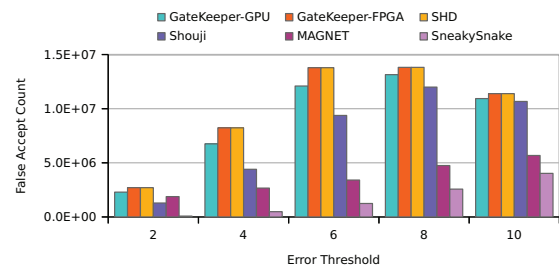

Fig. 1. False accept comparison with the other pre-alignment filters (GateKeeper-FPGA [2], SHD [3], MAGNET [4], Shouji [5], SneakySnake [6]) for candidate mappings produced by mrFAST with read length = 100bp.

IV. CONCLUSION

Compared to the original GateKeeper (FPGA) [2], GateKeeper-GPU is more accurate but it has a smaller acceleration on the entire read mapping process. On the other hand, it is a lot more promising for further improvements since it has a GPU code base. We believe that when a new short read mapper is constructed with a verification-aware hardware-software co-design, GateKeeper-GPU's impact on the entire mapping procedure can be more powerful than adapting it into an existing pipeline.

REFERENCES

[1] C. Alkan, J. M. Kidd, T. Marques-Bonet, G. Aksay, F. Antonacci, F. Hormozdiari, J. O. Kitzman, C. Baker, M. Malig, O. Mutlu *et al.*, "Personalized Copy Number and Segmental Duplication Maps Using Next-Generation Sequencing," *Nature Genetics*, vol. 41, no. 10, p. 1061, 2009.

[2] M. Alser, H. Hassan, H. Xin, O. Ergin, O. Mutlu, and C. Alkan, "Gate-Keeper: a New Hardware Architecture for Accelerating Pre-Alignment in DNA Short Read Mapping," *Bioinformatics*, vol. 33, pp. 3355–3363, Nov. 2017.

[3] H. Xin, J. Greth, J. Emmons, G. Pekhimenko, C. Kingsford, C. Alkan, and O. Mutlu, "Shifted Hamming Distance: a Fast and Accurate SIMD-Friendly Filter to Accelerate Alignment Verification in Read Mapping," *Bioinformatics*, vol. 31, no. 10, pp. 1553–1560, 2015.

[4] M. Alser, O. Mutlu, and C. Alkan, "MAGNET: Understanding and Improving the Accuracy of Genome Pre-Alignment Filtering," *arXiv preprint arXiv:1707.01631*, 2017.

[5] M. Alser, H. Hassan, A. Kumar, O. Mutlu, and C. Alkan, "Shouji: a Fast and Efficient Pre-Alignment Filter for Sequence Alignment," *Bioinformatics*, vol. 35, no. 21, pp. 4255–4263, 2019.

[6] M. Alser, T. Shahroodi, J. Gómez-Luna, C. Alkan, and O. Mutlu, "SneakySnake: a fast and accurate universal genome pre-alignment filter for CPUs, GPUs and FPGAs," *Bioinformatics*, 12 2020, btaa1015. [Online]. Available: https://doi.org/10.1093/bioinformatics/btaa1015