FA3C: FPGA-Accelerated Deep Reinforcement Learning

Björn Bebensee (2019–21343)

Topics in Artificial Intelligence

October 1, 2019

While traditional FPGA-based DNN accelerators focus mainly on improving inference speed by exploiting fixed-point arithmetic, Cho, Oh, Park, Jung and Lee [1] propose a new FPGA-based deep reinforcement learning platform which goes beyond inference to also improve training speed using single-precision floating-point arithmetic and which achieves higher performance than GPU-based platforms. The proposed platform is for the Asynchronous Advantage Actor-Critic (A3C) method, a state-of-the-art method in deep RL. With this approach the authors address the fact that the A3C algorithm's training process cannot be easily accelerated by GPUs. Besides speeding up the computation, FGPA-based platforms are more energy-efficient than GPU-based platforms.

In this paper Cho et al. first identify the major bottlenecks in A3C on a GPU-based platform and how these can be resolved using an FPGA-based platform. The main bottlenecks are limited off-chip data bandwidth as A3C generates large amounts of off-chip data traffic, small batch sizes in A3C and the overhead from frequent kernel launches by CUDA or OpenCL even for small computations. As the GPU-based approach does not fully utilize the GPU's computing power due to these bottlenecks, an FPGA-based platform can achieve higher performance and energy efficiency.

The proposed FPGA-based platform uses compute units (CUs) consisting of multiple processing elements (PEs) for faster DNN computation. Each CU can either perform an inference or a training task. They use a pair of CUs to execute inference and training tasks for the multiple agents but a higher number of pairs can be included in order to improve the throughput for all of the A3C agents. They test their design by training the A3C agents to play six different Atari 2600 games and compare performance and energy efficiency of the FPGA to conventional GPU-based platform. They find that FA3C achieves higher performance (for $n = 16$ agents the number of inferences per second (IPS) is 27.9% higher than A3C-cuDNN) while maintaining a 30% reduction in average power consumption compared to A3C-cuDNN, thus achieving a 1.62× better energy efficiency. The authors also show that FA3C achieves similar training trends and results as GPU-based platforms. Therefore a similar result can be achieved in less time due to the higher IPS.

References

[1] Cho, Hyungmin, et al. "FA3C: FPGA-Accelerated Deep Reinforcement Learning." *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*. ACM, 2019.