SPONSORED BY THE

Federal Ministry of Education and Research

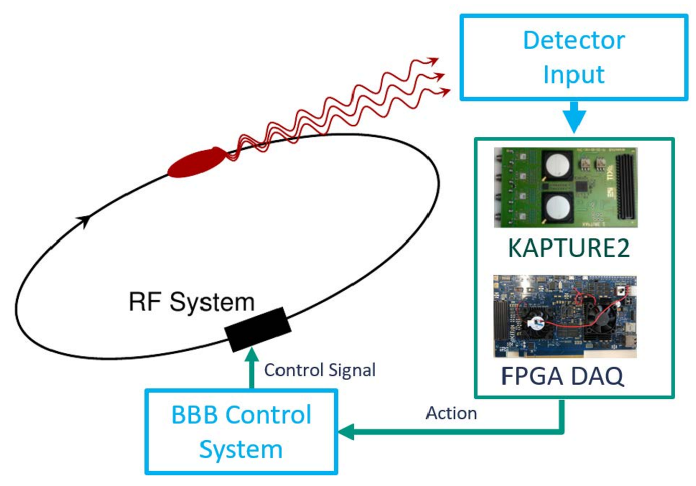
A Novel Readout Electronics and Reinforcement Learning Platform for Physics Experiments

W.Wang[1], O. Sander[1], M. Caselle[2], M. Patil[1], A. Ebersoldt[1], M. Balzer[1], M. Weber[1]

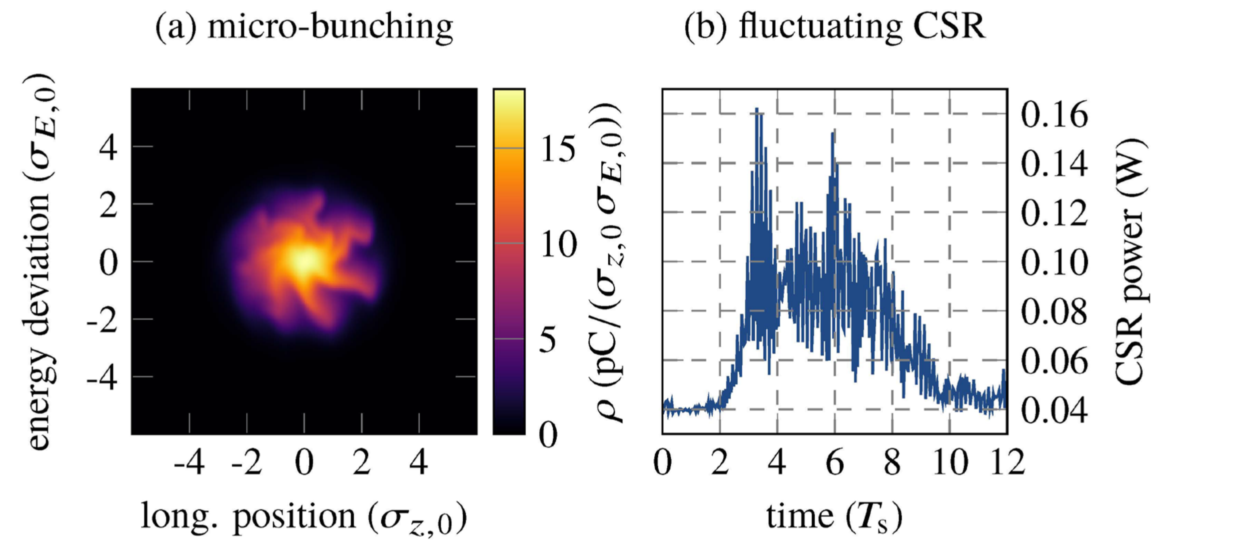
[1]Karlsruhe Institute of Technology [2]GSI, Darmstadt, Germany

Motivations

Modern beam diagnostics detectors generate large data volume of several gigabytes which needs to be processed in real time and latencies of microsecond order. In order to satisfy such requirements, we developed a novel heterogeneous system with ZYNQ Ultrascale+ based PCIe readout card and connected with GPU-based computing nodes. The KIT (Karlsruhe Institute of Technology) storage ring KARA (Karlsruhe Research Accelerator) aims at the stabilization of the coherent synchrotron radiation emitted in THz radiation to overcome the instability given by the instability. A reinforcement learning based method is planed to be implemented through a RF feedback system located at storage ring. IPE (Institute of Data Processing and Electronics) is developing a reinforcement learning hardware platform for the final realistic approach in this poster.



The hardware solution for RF control



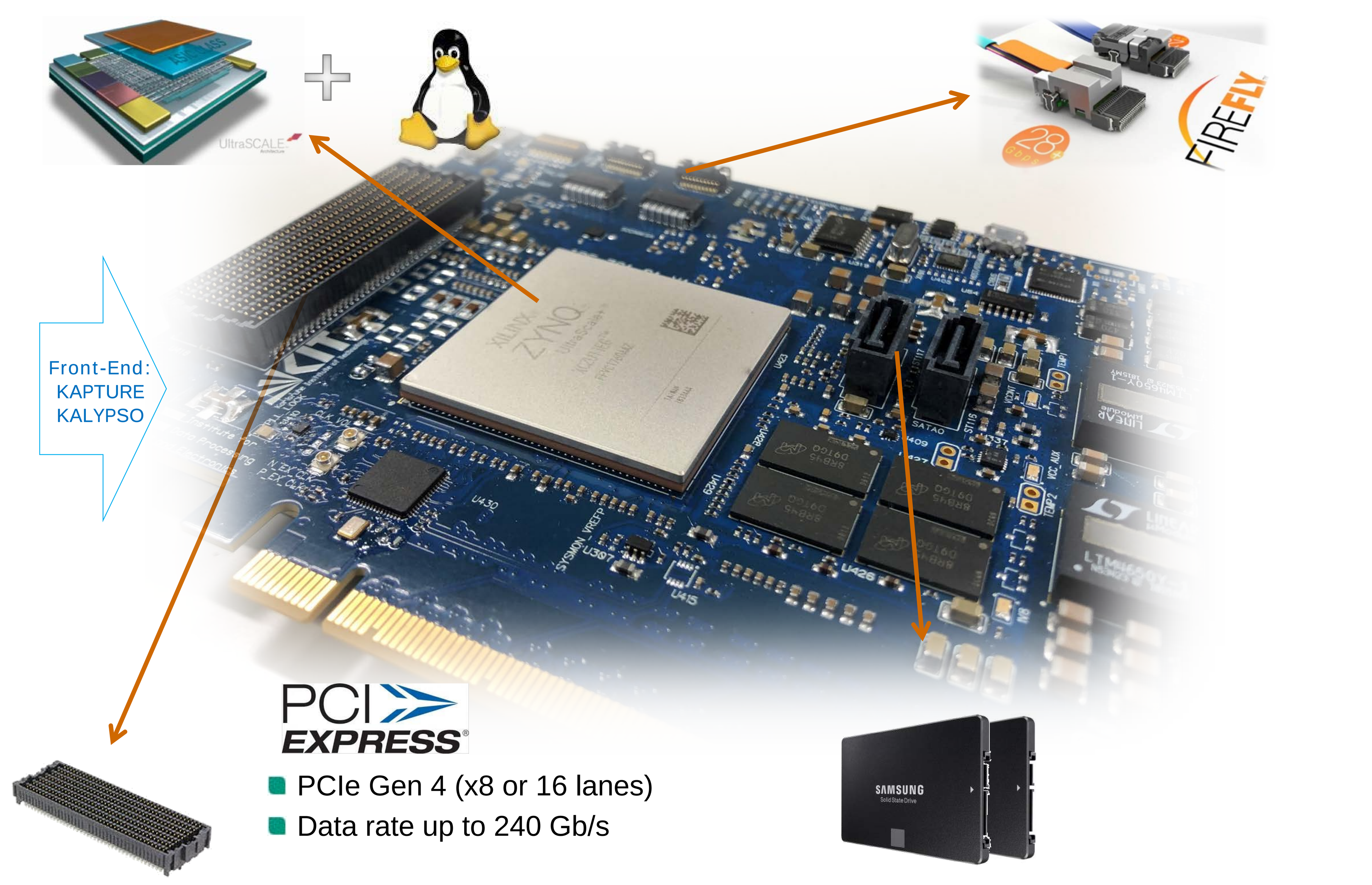
(a) The CSR self-interaction of the bunch causes the formation of micro-structures in the longitudinal phase space density.

(b) Their continuous variation leads to fluctuations in the emitted CSR power. The illustrated dynamics are simulated with the VFP solver Inovesa.

Boltz, Tobias, et al. "Feedback Design for Control of the Micro-Bunching Instability based on Reinforcement Learning." *10th Int. Partile Accelerator Conf.(IPAC'19), Melbourne, Australia, 19-24 May 2019.* JACOW Publishing, Geneva, Switzerland, 2019.

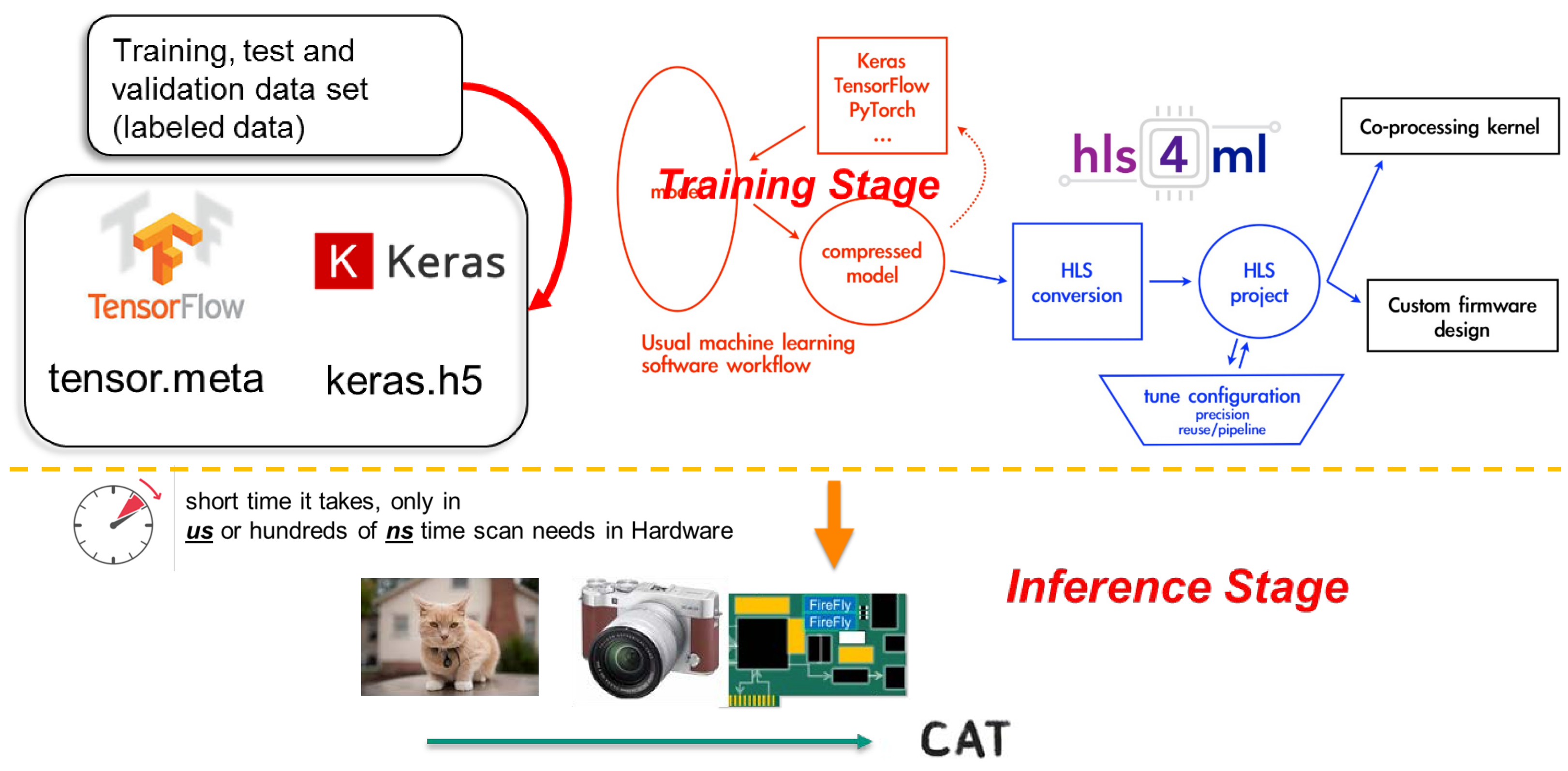
Novel PCIe ZYNQ Ultrascale+ Data Acquisition Board

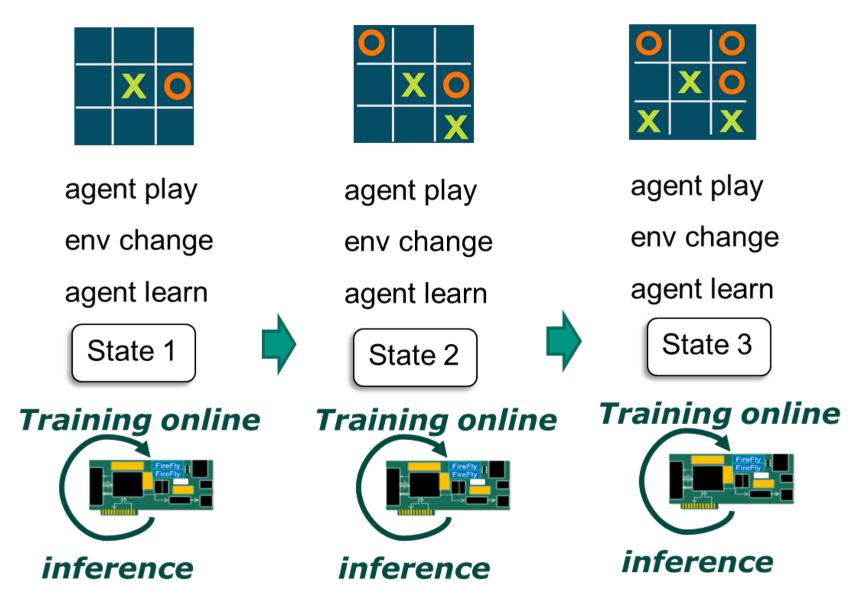
- Processor System (ARM): User Applications and Drivers based on Embedded Linux (eg. Yocto)
- Programmable Logic (FPGA): HDL IP Blocks
- 12 lanes @ 14Gb/s or 12 lanes @ 25 Gbps
- Routed on 12 GTY



- PCIe Gen 4 (x8 or 16 lanes)
- Data rate up to 240 Gb/s
- Extending the total bumber of Gigabit Transceivers to 32
- Compatible with standard FMC
- Two SATA connectors
- Data could be stored locally on SSD

The Reinforcement Learning and Supervised Learning Differences on Hardware

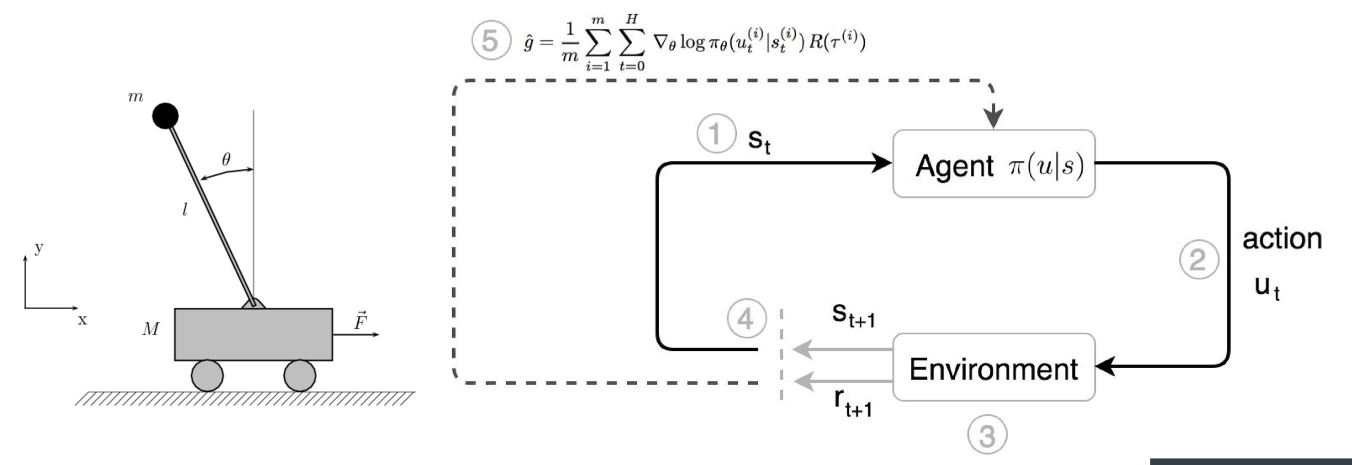


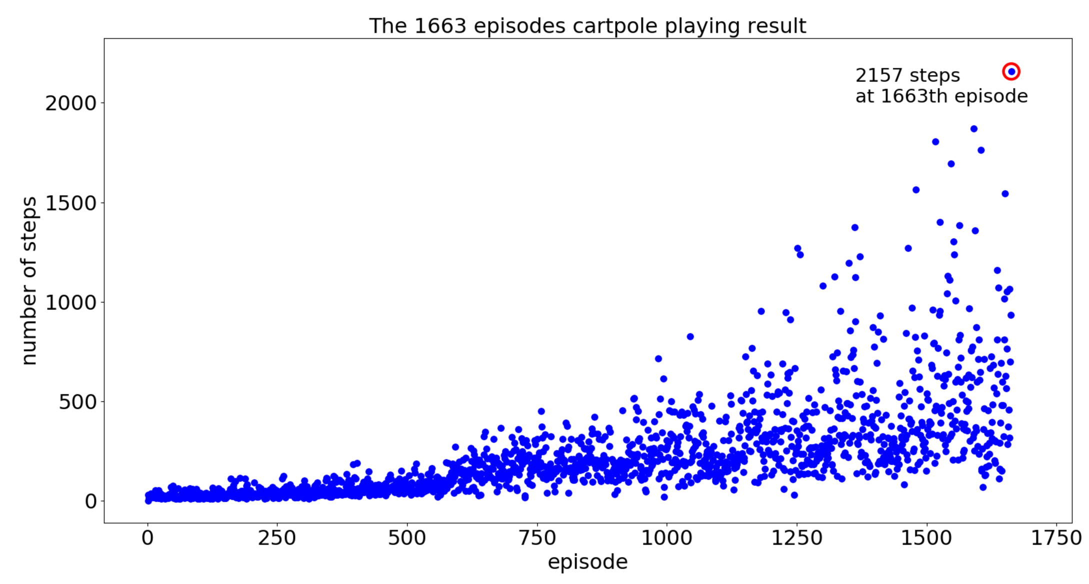


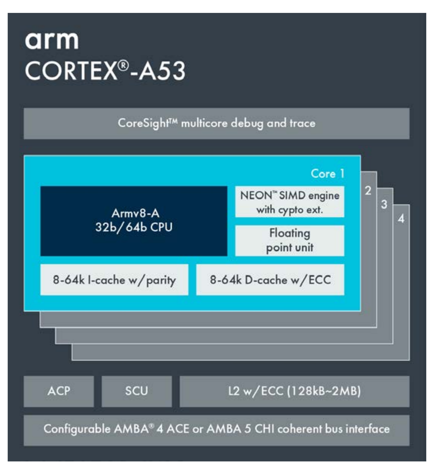
From the implementation point of view, the main difference between supervised learning (SL) and reinforcement learning (RL) is that the former can usually be done using an offline training process while for reinforcement learning, the learning process is continuous and has to happen during runtime in order to allow the agent to learn from past experience and allow an efficiently explore in its action space. This yields much higher timing demands than supervised learning.

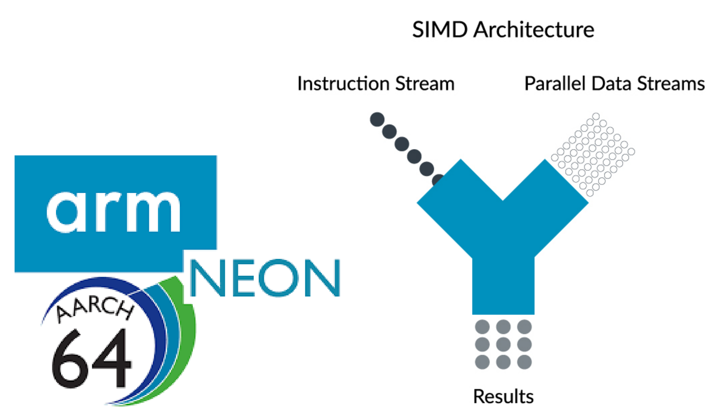
Two Implementation Result for RL on ARM Cortex-A53

The Policy Gradient Method on Hardware



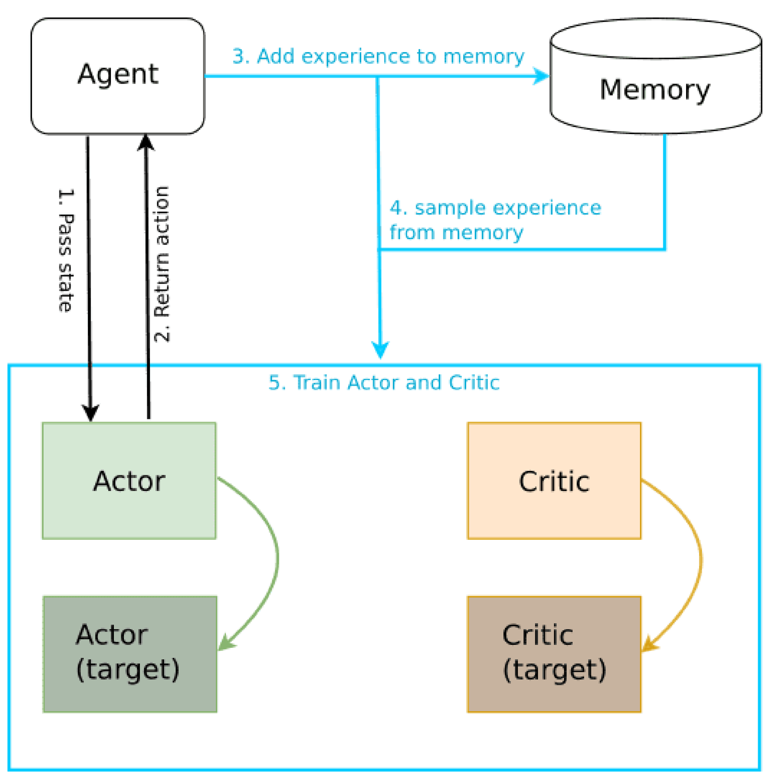




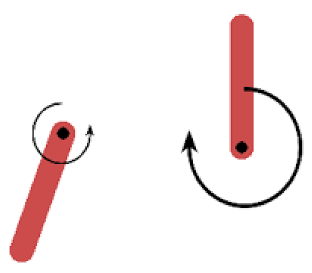


A full algorithm transform work is done on zynq. It shows the pole is kept balanced for 2157 steps at 1663th episode as shown in the red cycle at the top right of the picture. The whole game finished at the end of 1663th episode and took 161228.30 us, 193472022 clock cycles on ARM Cortex-A53.

The Deep Deterministic Policy Gradient Method on Hardware



A full algorithm transform work is done by C/C++ mix code. A near recent work is continue improves the speed by NEON intrinsic. It shows the a great improvement than Tensorflow on PC. A 32-batch of data on ARM only took around 13.35 us by average, based on this, the PC will take time range of few milli-seconds.



We first motivate the problem by demonstrating inadvertent generalization on the inverted pendulum task. The aim of the pendulum task is to train an agent to hold a pendulum mass steady above its pivot point by DDPG.

32-batch of data Training Time

TensorFlow on PC	Optimized C on PC	Optimized C on ARM
3145.00 us	91.00 us	13.36 us
3209.00 us	91.00 us	13.35 us
3100.99 us	92.00 us	13.36 us
3260.00 us	91.00 us	13.36 us
3472.00 us	91.00 us	13.37 us
2862.00 us	92.00 us	13.36 us
3338.00 us	91.00 us	13.36 us
2893.99 us	91.00 us	13.35 us
3480.99 us	91.00 us	13.36 us
3760.99 us	91.00 us	13.35 us
3167.00 us	92.00 us	13.36 us
4532.99 us	92.00 us	13.36 us
3822.00 us	91.00 us	13.36 us
2914.99 us	94.00 us	13.35 us
3111.99 us	91.00 us	13.36 us

The result shows that the full transport by C on ARM is much more faster than Tensorflow in the training 32-batch data. It is also shows a more stable time range (low variance) during 15 tests.

Contact Author: Weijia Wang
Email: weijia.wang@partner.kit.edu