MODEL-FREE CONTROL FOR DISTRIBUTED STREAM DATA PROCESSING USING DEEP REINFORCEMENT LEARNING

TENG LI
ZHIYUAN XU
JIAN TANG
YANZHI WANG
FROM SYRACUSE UNIVERSITY

PUBLISHED IN:
PROCEEDINGS OF THE VLDB ENDOWMENT 2018

INTRODUCTION

This paper focuses on general-purpose Distributed Stream Data Processing Systems (DSDPSs), which deal with processing of unbounded streams of continuous data at scale distributedly in *real or near-real time*. A fundamental problem in a DSDPS is the *scheduling problem* (i.e., assigning workload to workers/machines) with the objective of *minimizing average end-to-end tuple processing time*.

DISTRIBUTED SYSTEM

A distributed system is a network that consists of autonomous computers that are connected using a distribution middleware. They help in sharing different resources and capabilities to provide users with a single and integrated coherent network.

REINFORCEMENT LEARNING

An agent sees states $s^{(t)}$'s of an environment, takes actions $a^{(t)}$'s, and receives rewards $R^{(t)}$'s (or penalties)

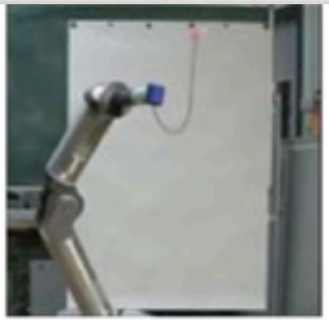
- Environment does not change over time
- The state of the environment may change due to an action
- Reward $R^{(t)}$ may depend on $s^{(t+1)}, s^{(t)}, \cdots$ or $a^{(t)}, a^{(t-1)}, \cdots$
- Goal: to learn the best policy $\pi^*(s^{(t)}) = a^{(t)}$ that maximizes the total reward $\sum_t R^{(t)}$

RL APPLICATIONS

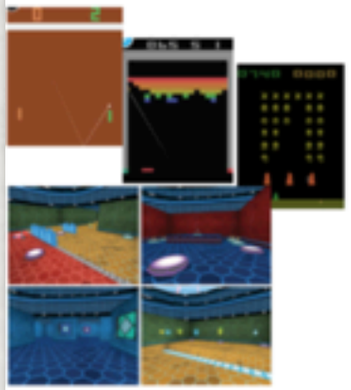
Kohl and Stone, 2004

Ng et al, 2004

Tedrake et al, 2005

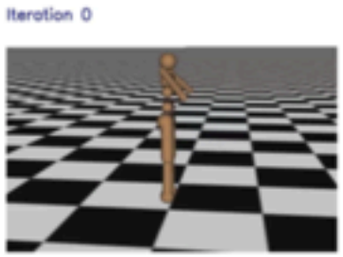


Kober and Peters, 2009



Mnih et al 2013 (DQN)
Mnih et al, 2015 (A3C)

Silver et al, 2014 (DPG)
Lillicrap et al, 2015 (DDPG)



Schulman et al,
2016 (TRPO + GAE)

Levine*, Finn*, et
al, 2016
(GPS)

Silver*, Huang*, et
al, 2016
(AlphaGo)

How to allocate workload in a large-scale distributed data processing system?

Goal: Minimize average end-to-end processing time

Traditional way:

- Distribute workload evenly over machines in the cluster in a round-robin matter.

However…

The communication delay is not considered.

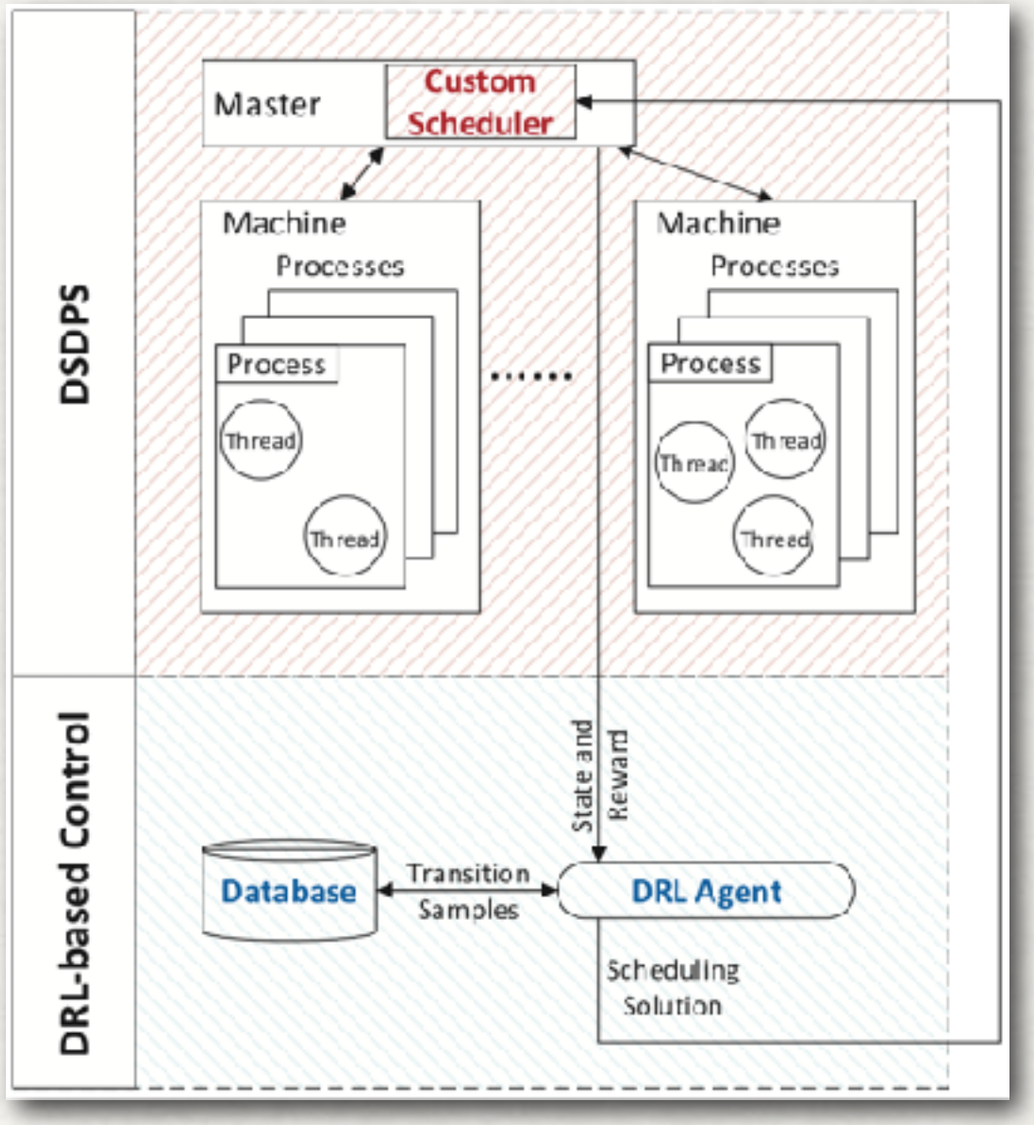
THE PROPOSED METHOD

Input:

- The current state of the distributed system.

Output:

- The distribution of the incoming workload.
- The rewards of the corresponding actions.



DISCUSSION

Premise:

- A large-scale distributed stream data processing system
- Unbounded streams of continuous data.
- Need to process data in real-time or near real-time.
- Average end-to-end tuple processing time should be considered.

Question:

- How to define states, actions and reward in the DRL model?
- Why does the DRL implementation surpass the traditional one?