Reinforcement Learning Software Ecosystems

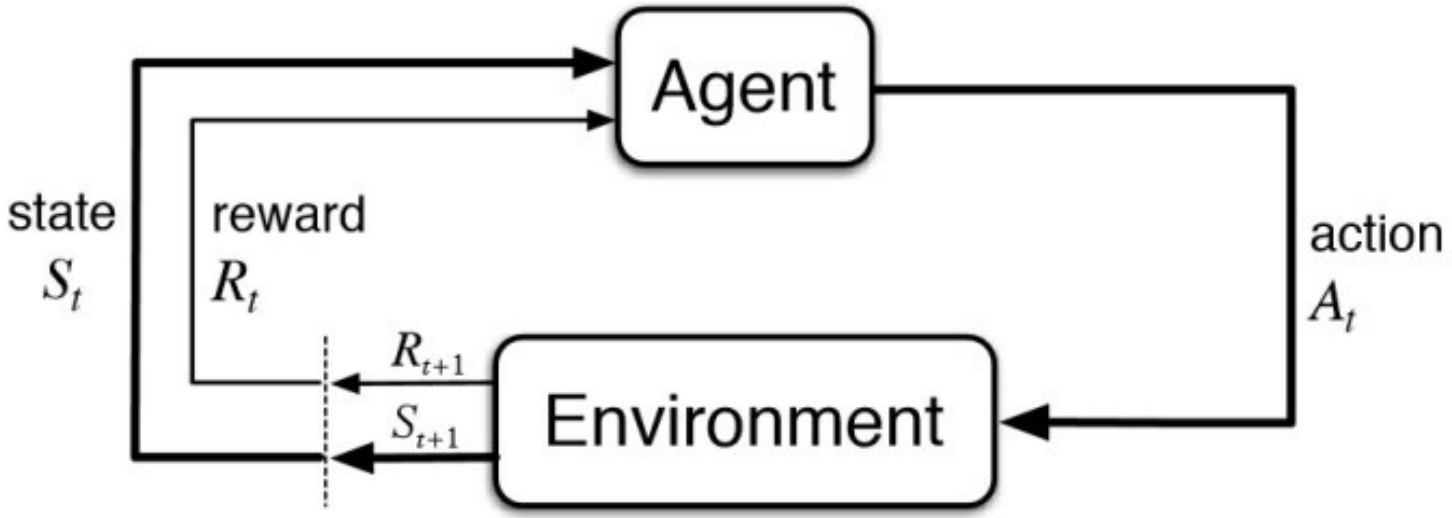
Eddie Rusu

Computational Scientist at LLNL

rusu1@llnl.gov

What is RL

Pick the *best* sequence of actions according to some objective

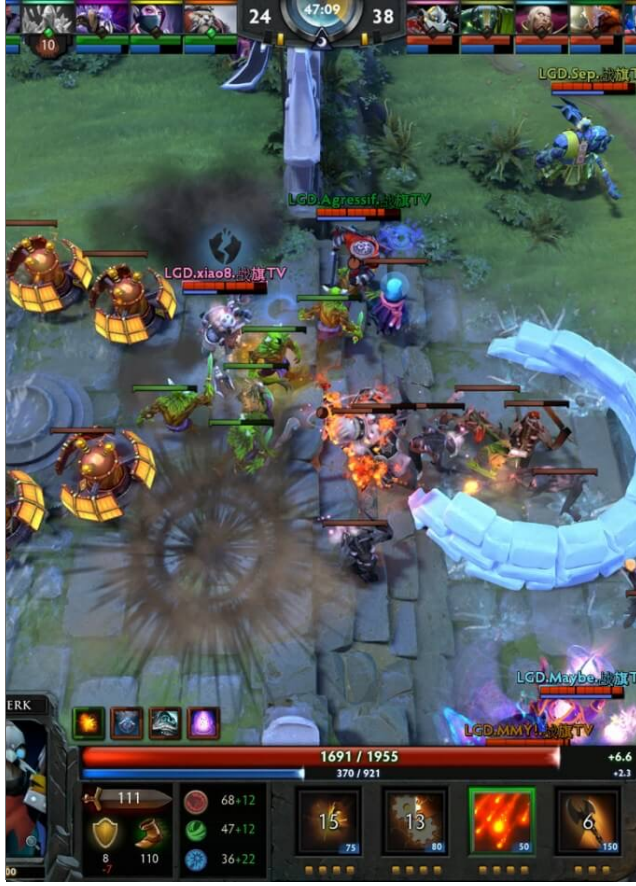


$$\{(S_0, A_0, R_0), (S_1, A_1, R_1), \dots, (S_n, A_n, R_n)\}$$

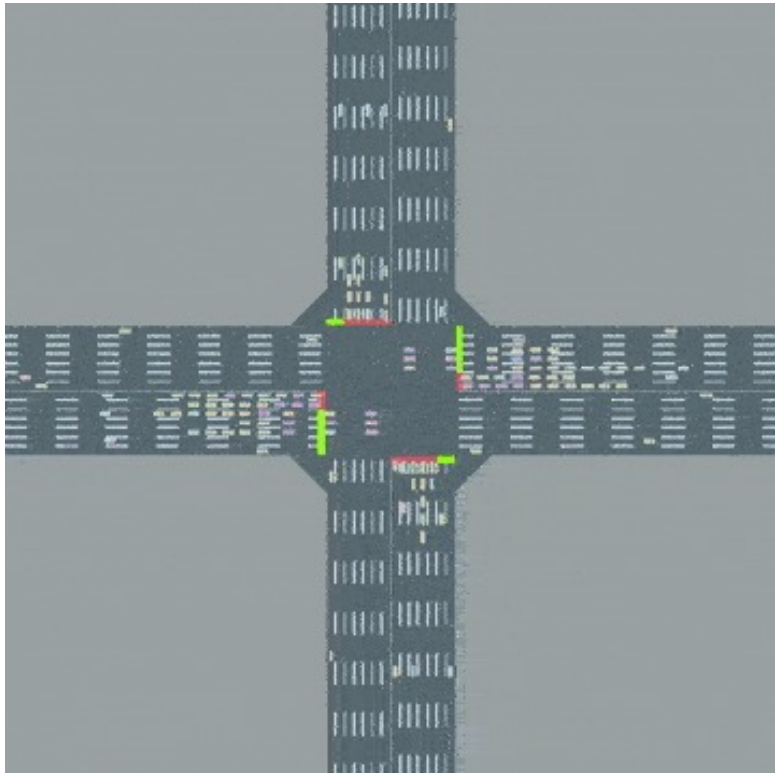
nature
THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
At last — a computer program that can beat a champion Go player PAGE 484
ALL SYSTEMS GO
SONGBIRDS A LA CARTE
SAFEGUARD TRANSPARENCY
WHEN GENES GOT 'SELFISH'

How research funders profit from hidden investments p. 1330
New books for budding scientists p. 1324
Drug leads for malaria pp. 1122 & 1129
Science
7 DECEMBER 2018
AAAS
A DIGITAL PRODIGY
AlphaZero teaches itself chess, shogi, and Go
pp. 1087, 1118, & 1140

nature
GAME PLAN
AI program learns to play StarCraft II to Grandmaster level


1691 / 1955
370 / 921

RL for games

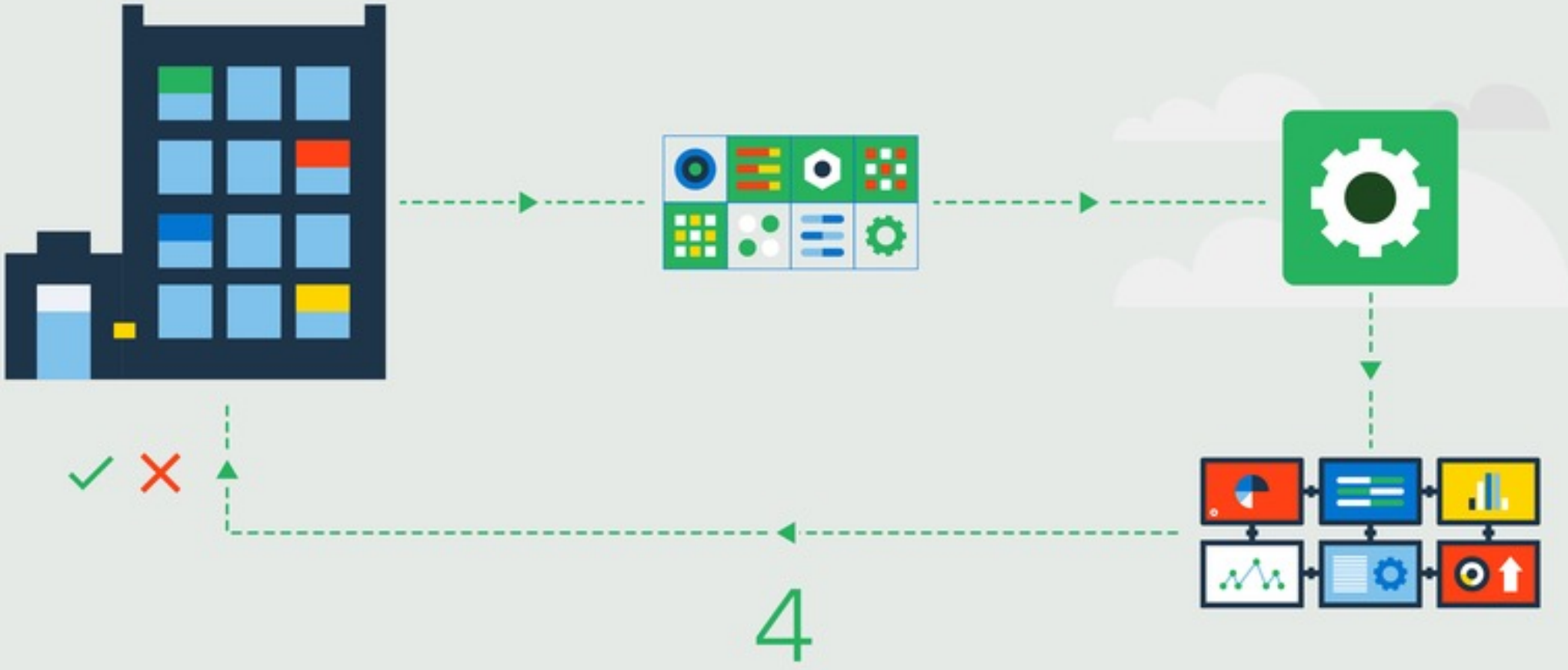


Transportation

Industrial Automation

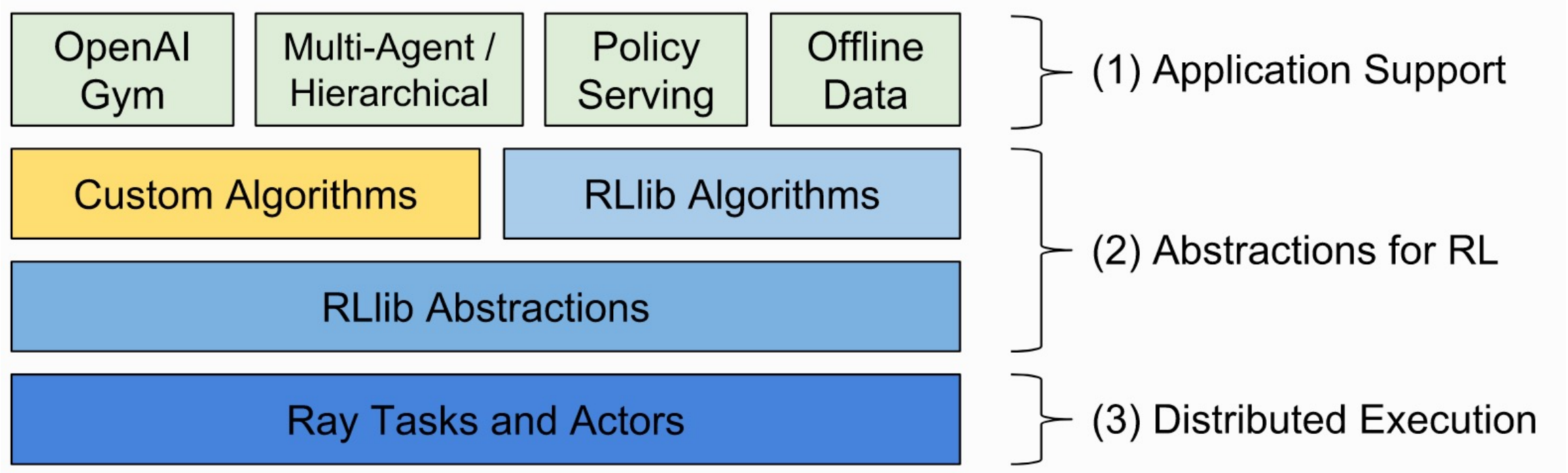
RL in Industry

Google Cooling Data Centers



Dopamine

Software Ecosystem



RLlib Design

RLlib Algorithms

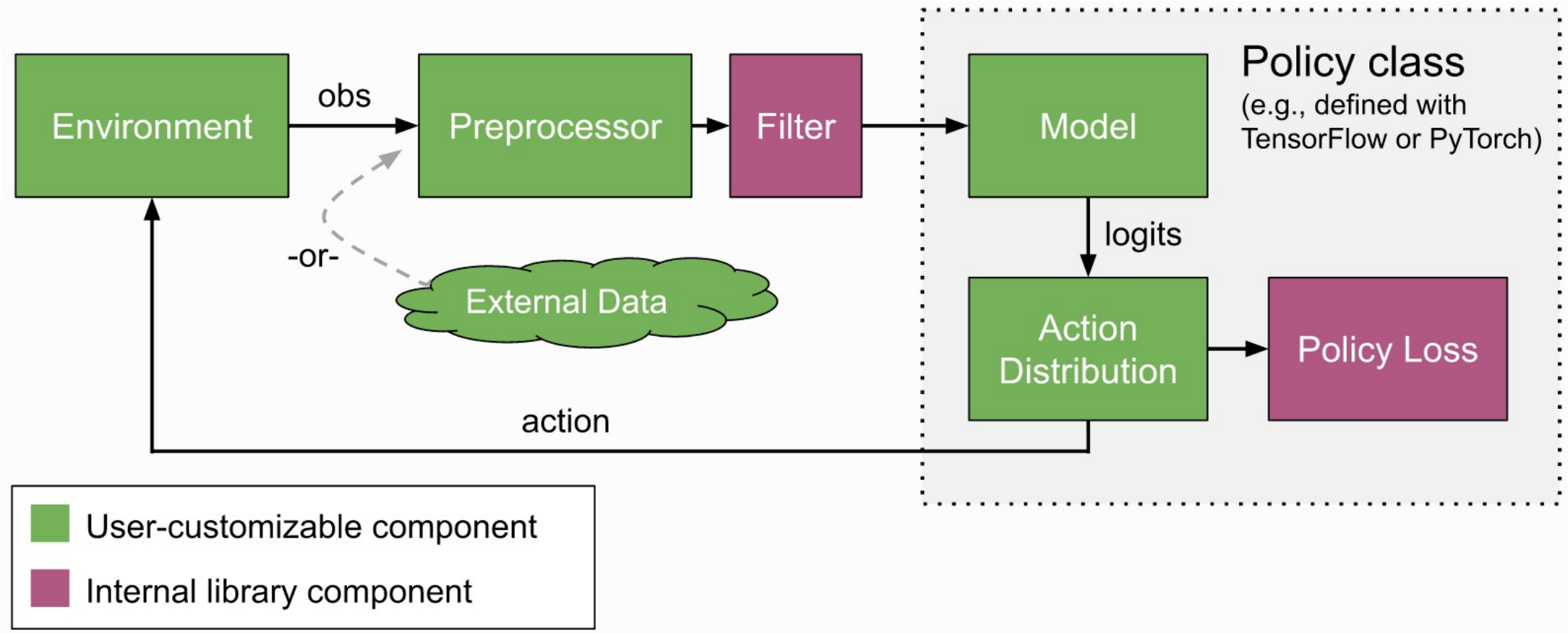
Single Agent State of the Art Algorithms

Algorithm	**Model Support (+DNN, +CNN)**	**Parallel Support (+CPU)**
A2C, A3C	RNN, Attention, Auto-regression	GPU
Behavior Cloning (BC)	RNN	
DDPG, TD3		GPU
DQN, Rainbow		GPU
IMPALA	RNN, Attention, Auto-regression	GPU
Imitation Learning (MARWIL)	RNN	GPU
PG	RNN, Attention, Auto-regression	GPU
PPO, APPO	RNN, Attention, Auto-regression	GPU
R2D2	RNN, Attention, Auto-regression	GPU
SAC		GPU
LinUCB, LinTS		

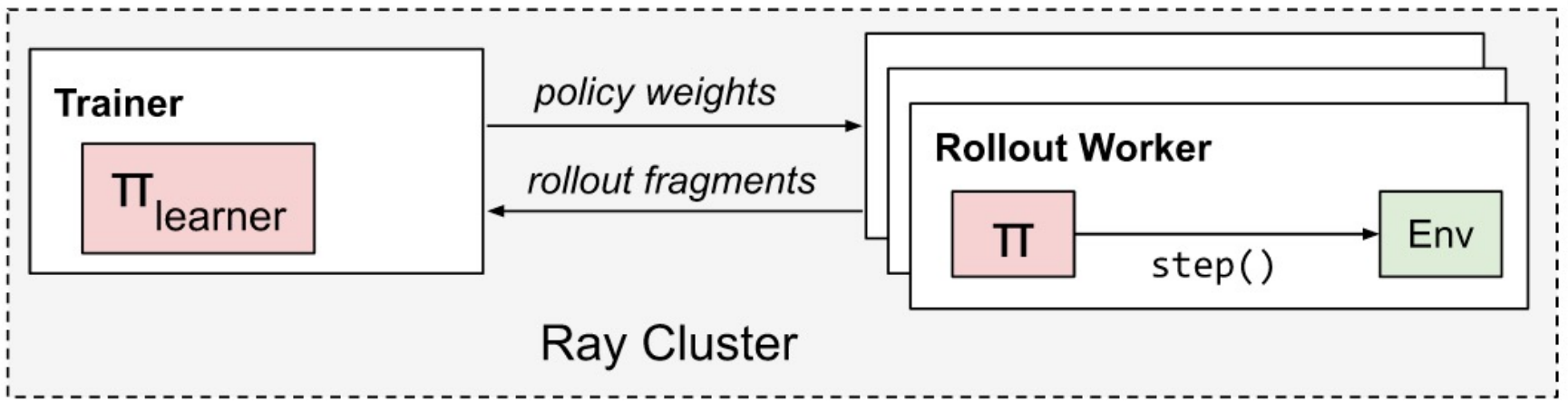
Multiagent Algorithms

Algorithm
QMIX (QMIX, VDN, IQN)
MADDPG
Parameter Sharing*
Fully Independent Learning*
Shared Critic Methods*

* Use any Algorithm in MARL



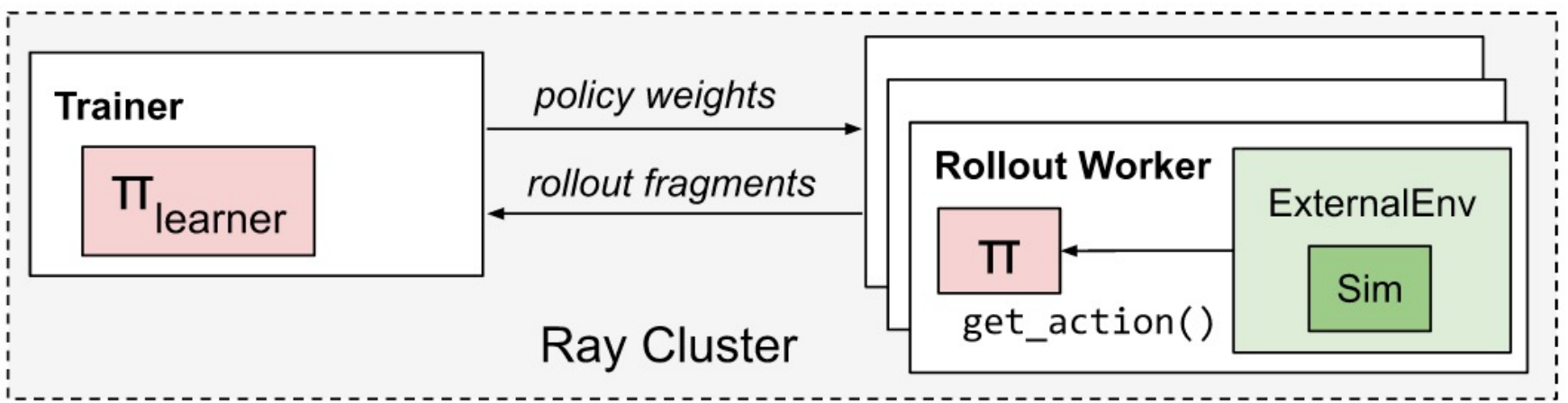
RLlib Customizability



(1) Standard environments (e.g., gym.Env, MultiAgentEnv types) are created and stepped by RLlib rollout workers.

RLlib Application Support

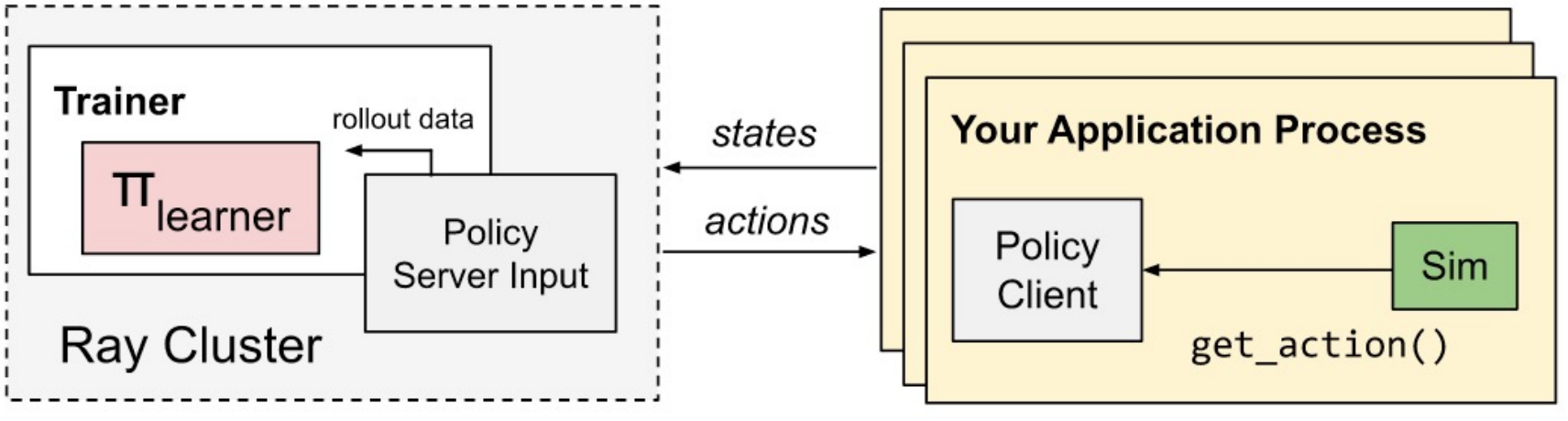
- Standard “toy” environments in gym.Env and MultiAgentEnv



(2) External environments (ExternalEnv) run in their own thread and pull actions as needed. RLlib still creates one external env class instance per rollout worker.

RLlib Application Support

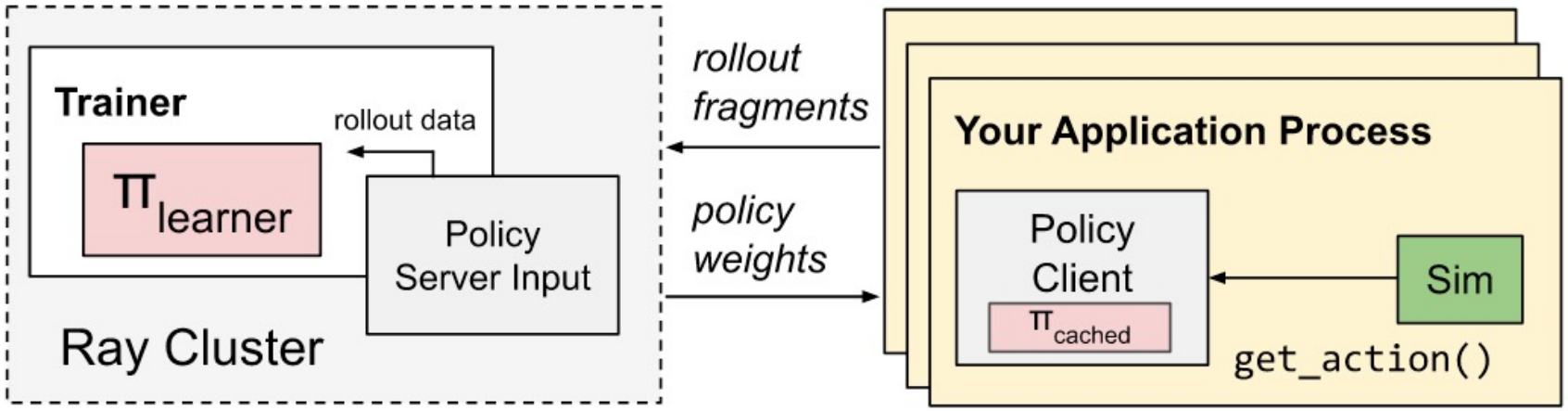
- Standard “toy” environments in gym.Env and MultiAgentEnv
- External environment controls learning



(3) Applications running outside the Ray cluster entirely can connect to RLlib using PolicyClient, which computes actions remotely over RPC.

RLlib Application Support

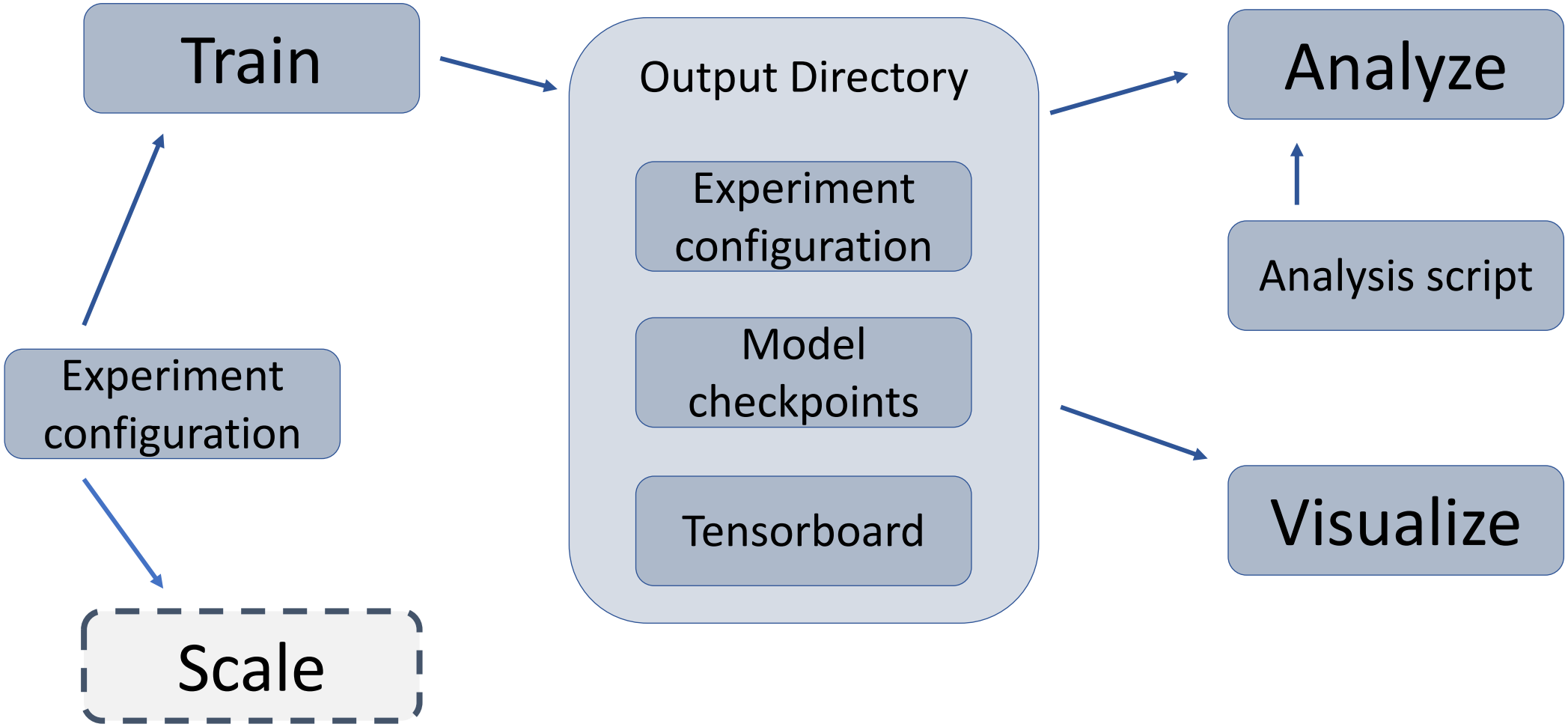
- Standard “toy” environments in gym.Env and MultiAgentEnv
- External environment controls learning
- Compute actions over RPC



(4) PolicyClient can be configured to perform inference locally using a cached copy of the policy, improving rollout performance.

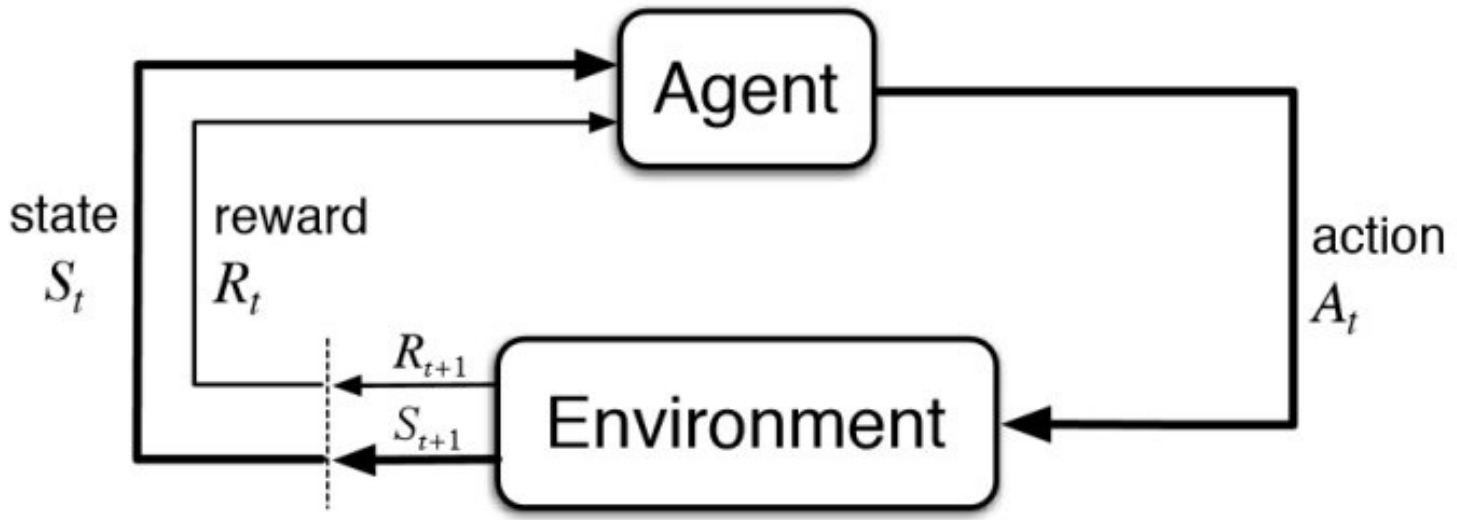
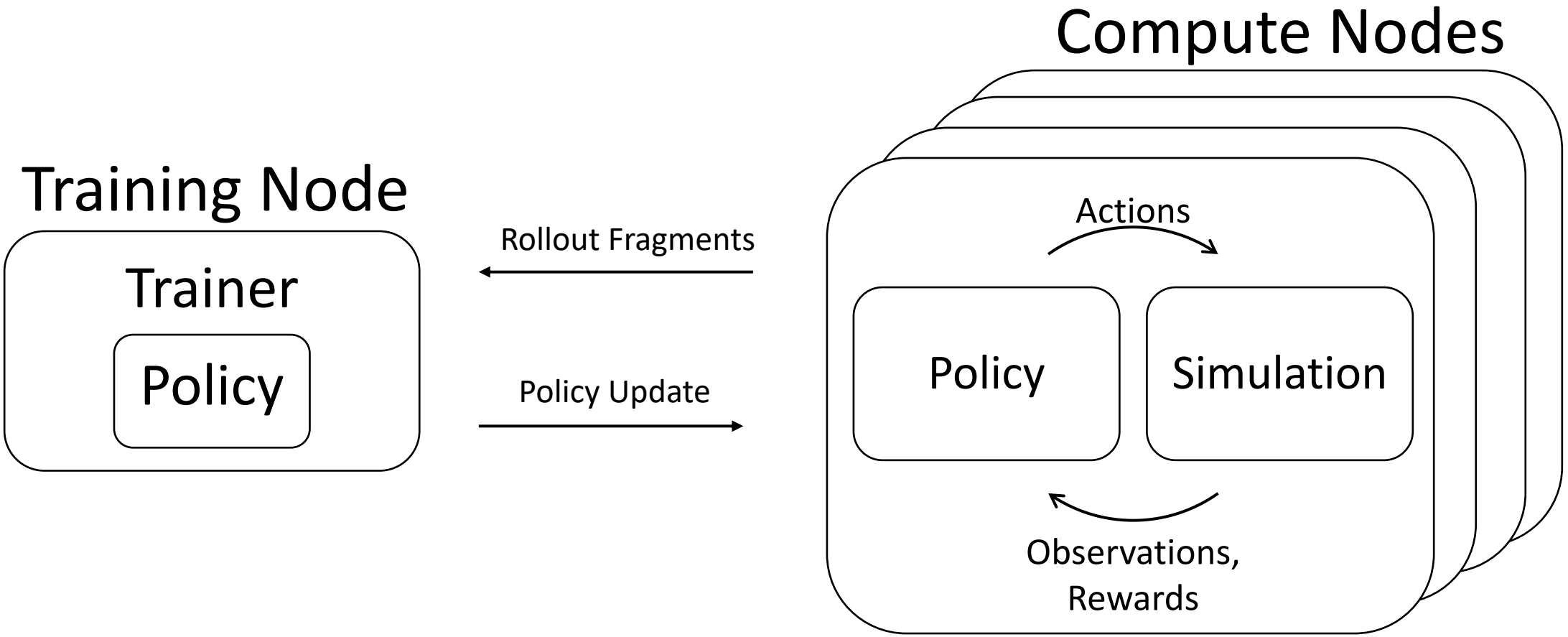
RLlib Application Support

- Standard “toy” environments in gym.Env and MultiAgentEnv
- External environment controls learning
- Compute actions over RPC
- Copy policy to external application



Abmarl
Three scripts in command line interface

Reinforcement Learning and Software Ecosystems



Reinforcement Learning and Software Ecosystems