

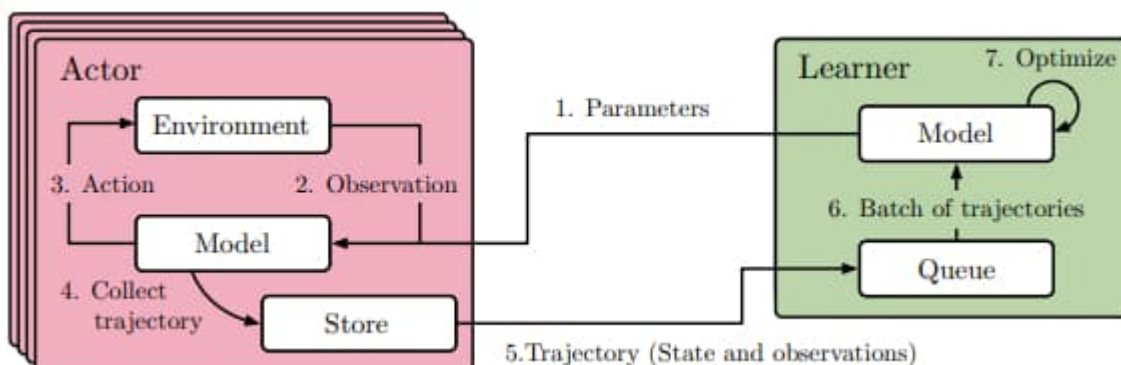
IA : Google s'attaque à la question des coûts

Comment rendre moins coûteux l'apprentissage par renforcement ? En permettant aux algorithmes d'exploiter plus efficacement les ressources à leur disposition.

Google aborde la problématique à travers le projet [SEED RL](#), dont il vient d'officialiser la mise en *open source*.

À la base de ces travaux, un modèle (ou agent) d'apprentissage par renforcement que sa filiale DeepMind avait dévoilé voilà deux ans : [IMPALA](#).

Ce dernier repose sur une architecture dite « acteur-critique ».



Divisé en une multitude de sous-agents déployés en parallèle sur des milliers de machines, l'acteur a pour rôle d'explorer un environnement et d'en tirer des expériences à partir d'une politique. Ces séquences d'états, d'actions et de récompenses, il les transmet à un apprenant centralisé (le critique) qui traite les expériences et améliore la politique.

Inférence déportée

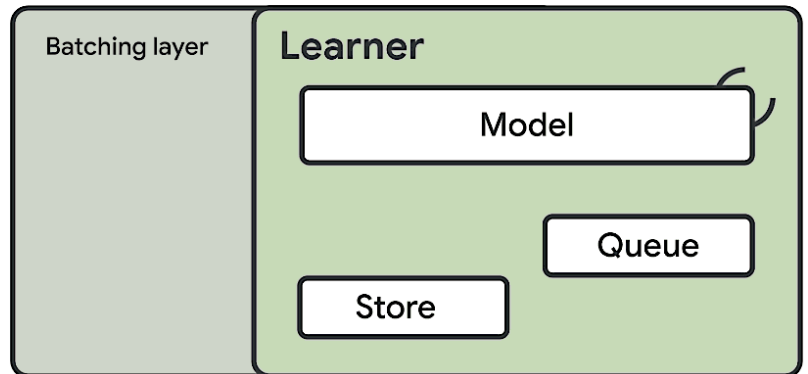
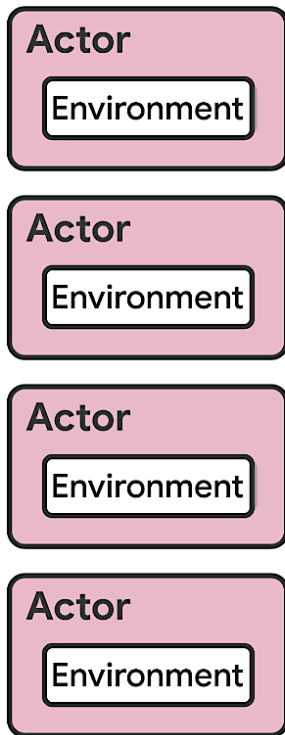
Traditionnellement, on fait tourner les acteurs sur des CPU. Les critiques, moins nombreux, s'exécutent sur des GPU.

Une telle architecture présente, d'après Google, des inconvénients :

- L'inférence est réalisée sur CPU, ce qui nécessite d'augmenter rapidement le nombre d'acteurs à mesure qu'un modèle croît.
- La communication avec les critiques consomme de la bande passante.
- Combiner exploration de l'environnement et inférence sur une même machine n'est pas favorable à un usage optimisé des ressources.

[SEED RL](#) remodèle cette architecture avec deux points principaux :

- L'inférence se fait côté critique(s).
- Pour les transferts de données, le *framework* [gRPC](#) est mis à contribution. Performance annoncée : jusqu'à un million de requêtes par seconde vers chaque machine.



Comme IMPALA, SEED RL exploite V-trace. Cet algorithme de correction hors politique permet de normaliser l'apprentissage des agents pour éviter la décohérence temporelle.

Sur l'environnement [DeepMind Lab](#), Google annonce 2,4 millions de trames traitées par seconde avec 64 cœurs TPU (du nom de ses puces dédiées à l'IA). IMPALA requiert trois à quatre fois plus de ressources pour atteindre la même vitesse.

Photo d'illustration © [mikemacmarketing](#) via [VisualHunt](#) / [CC BY](#)