



**TRƯỜNG ĐẠI HỌC
BÁCH KHOA HÀ NỘI**
HANOI UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Mô phỏng thuật toán Reinforcement Learning cho Serverless

Future Internet Laboratory

Nguyễn Phạm Trung Hiếu

Ngày 5 tháng 5 năm 2024

ONE LOVE. ONE FUTURE.

- ▶ Nội dung cần chỉnh sửa theo yêu cầu bài toán

 - Nội dung đã hoàn thành

 - Vấn đề còn lại

- ▶ Các lưu ý cần nhớ và phát triển

- ▶ Kết luận

- ▶ Nội dung cần chỉnh sửa theo yêu cầu bài toán

 - Nội dung đã hoàn thành

 - Vấn đề còn lại

- ▶ Các lưu ý cần nhớ và phát triển

- ▶ Kết luận

- ▶ Nội dung cần chỉnh sửa theo yêu cầu bài toán

 - Nội dung đã hoàn thành

 - Vấn đề còn lại

- ▶ Các lưu ý cần nhớ và phát triển

- ▶ Kết luận

Nội dung đã hoàn thành

- Cập nhật thời gian tính reward và đưa ra action thành 10s.
 - Sửa hàm `_get_request()` có chức năng nhận request theo phân phối Poisson thay vì nhận một lượng request cố định.
 - Tạm thời sử dụng công thức tính reward đơn giản = số request đến WarmCPU.
 - Thêm điều kiện kiểm tra xem container nào đang chuyển trạng thái hoặc không:
 - container đang chuyển thì không đưa ra action → chọn action khác.
 - container không chuyển thì đưa ra action bình thường.
- Thêm ma trận `_is_changing_matrix` với các phần tử 0 là không chuyển trạng thái và 1 là đang chuyển trạng thái.

- ▶ Nội dung cần chỉnh sửa theo yêu cầu bài toán

 - Nội dung đã hoàn thành

 - Vấn đề còn lại

- ▶ Các lưu ý cần nhớ và phát triển

- ▶ Kết luận

- Sửa hàm `_get_reward()` tạm thời bỏ qua công thức tính reward phức tạp, đặt thành:

$$Reward = n_{request-to-L2} - \alpha \times t_{pending} - \beta \times r$$

trong đó:

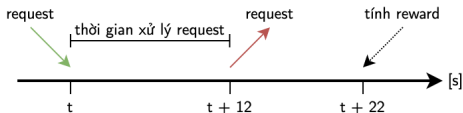
- $n_{request-to-L2}$ là số request đến WarmCPU là số request được xử lý xong trong khoảng thời gian 10s.
 - $t_{pending}$ là thời gian chờ là thời gian chuyển trạng thái sang WarmCPU để xử lý request (lưu ý 2-3 trường hợp).
 - r là tài nguyên sử dụng để đánh giá (RAM, CPU...).
- Đồng bộ về mặt thời gian xử lý song song \rightarrow Ước tính được request có được xử lý xong hay không, rồi mới đợi một khoảng chờ cho đến khi tính reward.

Vấn đề còn lại: Thời gian xử lý song song

Ví dụ, một request được xử lý trong 12s, mô phỏng yêu cầu cứ 10s tính reward một lần:

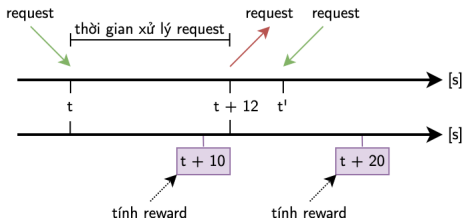
Mô phỏng hiện tại:

✗ Chờ xong mới xử lý



Mô phỏng yêu cầu:

✓ Xử lý song song



► Nội dung cần chỉnh sửa theo yêu cầu bài toán

Nội dung đã hoàn thành

Vấn đề còn lại

► Các lưu ý cần nhớ và phát triển

► Kết luận

- RAM và tài nguyên sử dụng trong khi chuyển trạng thái không ảnh hưởng.
- Giả sử hệ thống ban đầu toàn bộ đều ở WarmDisk (thay vì Null như trước).
- Công thức tính reward ban đầu:

$$Reward = a \times Profit - b \times Energy_{cost} - Penalty_{delay} - Penalty_{abandone}$$

- Tìm hiểu và áp dụng phương pháp tính reward theo cập nhật mềm (long term).
- Cần tăng λ lên khoảng 500 hoặc hơn \rightarrow đảm bảo hệ thống không quá rảnh.

- ▶ Nội dung cần chỉnh sửa theo yêu cầu bài toán

 - Nội dung đã hoàn thành

 - Vấn đề còn lại

- ▶ Các lưu ý cần nhớ và phát triển

- ▶ Kết luận

- Link Github mô phỏng:
[*https://github.com/owofuyuki/reinforcement-learning-for-serverless*](https://github.com/owofuyuki/reinforcement-learning-for-serverless)
- Kết quả mô phỏng chưa đánh giá được vấn đề cho bài toán.
- Chỉnh sửa mô phỏng...

**CẢM ƠN MỌI NGƯỜI
ĐÃ LẮNG NGHE!**

