Remover marca d'água agora

UDESC UNIVERSIDADE DO ESTADO DE SANTA CATARINA

Seminário de Iniciação Científica
Universidade do Estado de Santa Catarina
29º SIC UDESC

ANÁLISE DOS MODELOS DE APRENDIZADO DE MÁQUINA POO E DQN APLICADO A NUVEM COMPUTACIONAL

Renato Tanaka[1], Maurício A. Pillon[2]
[1] Acadêmico(a) do Curso de Ciência da Computação CCT - bolsista PROBIC/UDESC
[2] Orientador, Departamento de Ciência da Computação – CCT – mauricio.pillon@udesc.br.

Palavras-chave: *Aprendizado de Máquina, Nuvens Computacionais, Escalonamento e Energia.*

Introdução

O fornecimento de recursos computacionais sob demanda, implementado através de plataformas de nuvens computacionais, tem movimentado pesquisas científicas e novos negócios na indústria nesta última década. O modelo de serviço de nuvem computacional IaaS (*Infrastructure as a Service*), entrega a gerência integral de recursos virtuais (e.g. vCPU) aos seus inquilinos. Um *Cloud Operating System* (COS) aloca Máquinas Virtuais (MV) solicitadas por inquilinos em um servidor físico de acordo com a política de alocação da nuvem, implementada através de algoritmos de escalonamento de MVs. Aos algoritmos de escalonamento de MVs, associados a COS, cabe a difícil tarefa de escolha de alocação do melhor servidor físico compatível com a solicitação do inquilino. A escolha deve estar de acordo com a política de distribuição de recursos da nuvem, porém ele não tem conhecimento prévio do uso efetivo de recursos da aplicação que executa dentro da MV. *Overbooking* caracteriza-se pela prática de alocação de MVs, em servidores físicos, com capacidade de recursos físicos inferior a soma dos recursos virtuais solicitados pelo conjunto de MVs hospedadas no servidor. Se por um lado, o uso das técnicas favorece a redução de energia, por outro, prejudica a disponibilidade e qualidade dos serviços.

As políticas de alocação de recursos equalizam métricas como disponibilidade, consolidação e consumo de energia. Neste contexto, o presente trabalho visa analisar a capacidade de aprendizado das técnicas *Deep Q-Network* (DQN) e *Proximal Policy Optimization* (PPO) associadas ao problema de alocação de MVs em um ambiente de nuvens computacionais. Um algoritmo de escalonamento de MVs baseado em aprendizado de máquina é dependente de uma fórmula de recompensa, a qual impacta diretamente nos objetivos do algoritmo (e.g. consolidação, economia energética, disponibilidade). Para a análise, definiu-se as métricas e elaborou-se cinco modelos de recompensas distintos para serem aplicados em ambos as técnicas. Nos resultados preliminares, PPO obteve scores superiores a DQN em todas as recompensas elaboradas.

Análise Experimental

O conjunto de recompensas analisadas (Tabela 1) considera a quantidade de recursos, a capacidade de processamento, limiares máximos de *overbooking* dos servidores e de fragmentação do *datacenter* (DC). A técnica de aprendizado de máquina tem como objetivo maximizar a recompensa que se encontra em uma escala entre 0 e 1. As recompensas são formalizadas na Tabela 1, cujo, α representa a quantidade de recursos de CPU; β a capacidade total de CPU; $limiar_{overbooking}$, o limite máximo de *overbooking* aceito; $reward_{frag}$, a recompensa de acordo com o nível de consolidação do servidor, $servs_{ativos}$, conjunto de servidores

Remover marca d'água agora

com capacidade de aceitar a requisição sem *overbooking*; e $servs_{semOB}$, servidores sem sobrecarga.

	Fórmulas
R1	$reward = max\left(0, \quad 1 - \frac{\alpha}{limiar_{overbooking} * \beta} + reward_{frag}\right)$
R2	$reward = max\left(0, \quad 1 - \frac{\alpha}{limiar_{overbooking} * \beta}\right)$
R3	$reward = 1 - \frac{servs_{ativos}}{servs_{semOB}}$
R4	$reward = -5 * (\frac{taxa_{OB} - 1}{5})^2 + 1$
R5	$reward = \frac{1}{serv_{rank}}$

Tabela 1: *Fórmulas das recompensas R1 à R5.*

Os resultados apresentados na Figura 1 foram obtidos através de simulação. A métrica de comparação foi o score, tradicionalmente utilizado na área de aprendizado de máquina para quantificar o grau de aprendizado da técnica em função do tempo ou da carga. As técnicas analisadas foram PPO e DQN com a aplicação de cargas sintéticas controladas. A infraestrutura é constituída de 256 servidores e a carga conta com 1000 requisições de MV iguais.

pdfelement

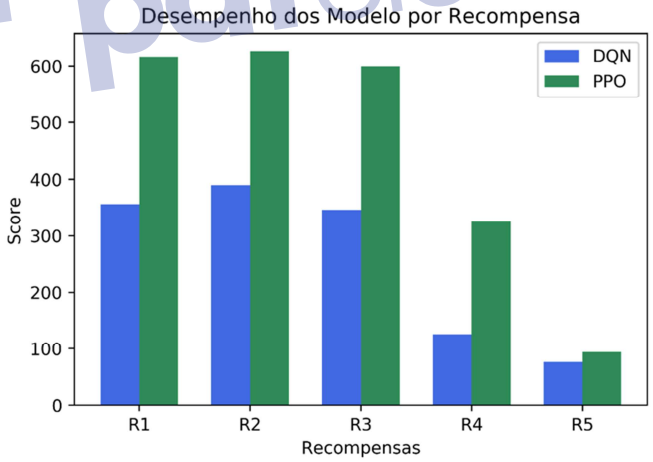

Fig.1: *PPO vs DQN: Análise de score de acordo com as recompensas*

PPO obteve melhores resultados do que DQN em todos cenários, atingindo um ganho de 25% com R3, em relação ao *score* máximo atingível. Como trabalhos futuros, faz-se necessário a aplicação da técnica de PPO, que se mostrou mais adequada, no escalonamento de recurso em nuvem.

Referências:

CHENG, Mingxi; LI, Ji; NAZARIAN, Shahin. DRL-cloud: Deep reinforcement learning-based resource provisioning and task scheduling for cloud service providers. **23rd Asia and South Pacific Design Automation Conference (ASP-DAC)**, Jeju Island , Coréia do Sul, 25 jan. 2018.