



Heurísticas Aumentadas Baseadas em Preditores de Demanda para Planejamento de Capacidade em Plataformas de IaaS

Moab Alves Silva ¹, Fábio Jorge Almeida Morais ²

RESUMO

As plataformas de Computação em Nuvem no modelo IaaS oferecem diferentes mercados para a aquisição de instâncias de máquinas virtuais. No mercado sob demanda o usuário adquire uma instância, pagando apenas pelo tempo utilizado, podendo cancelar a aquisição quando ela não for mais necessária. Enquanto no mercado de reserva, um usuário reserva uma instância por um período maior (1 ou 3 anos na AWS) se comprometendo a pagar por ela durante esse período. Apesar de ser um comprometimento a longo prazo, o mercado de reserva oferece descontos no custo incorrido por hora em comparação ao mercado sob demanda, o que torna ele mais vantajoso para demandas de longo prazo. Todavia, a decisão sobre qual mercado escolher não é uma tarefa trivial, uma vez que envolve tanto o prazo de encerramento de contratos existentes quanto as demandas futuras a serem supridas. Para mitigar esse problema, a literatura sugere o uso de heurísticas que consideram a demanda atual e histórico de compras para a tomada de decisão. No entanto, embora ofereçam um desempenho mínimo garantido, estas são limitadas pela falta de conhecimento sobre a demanda futura, que influencia significativamente a qualidade das decisões no planejamento de capacidade. Este trabalho investiga como a integração de modelos de predição de séries temporais com heurísticas online pode melhorar o planejamento de capacidade reduzindo os custos incorridos pela alocação de recursos. Os resultados mostram que é possível obter reduções no custo da alocação usando predições de curto ou longo prazo geradas a partir de modelos de predição de séries temporais presentes na literatura.

Palavras-chave: Computação em Nuvem, Planejamento de Capacidade, Heurísticas Online e Aumentadas.

¹ Aluno do curso de Ciência da Computação, Departamento de Sistemas e Computação, UFCG, Campina Grande, PB, e-mail: moab.alves.silva@ccc.ufcg.edu.br

² Doutor, Professor, Departamento de Sistemas e Computação, UFCG, Campina Grande, PB, e-mail: fabio@computacao.ufcg.edu.br

Heurísticas Aumentadas Baseadas em Preditores de Demanda para Planejamento de Capacidade em Plataformas de IaaS

ABSTRACT

IaaS cloud computing providers offer different pricing models for acquiring virtual machine instances. In the on-demand pricing model, users acquire instances and pay only for the time usage, with the flexibility to cancel the acquisition whenever it is no longer needed. Differently, the reserved model offers instances for long periods (1 or 3 years on AWS), where the users are committing to pay for the whole reserve period. Despite being a long-term commitment, the reserved model has cost reductions compared to hourly acquisition costs in the on-demand model. Thus, the reserved model may save costs for long-term demands, unlike the on-demand model indicated for short-term and varied demands. However, deciding the pricing model to use is not a trivial task, as it involves contracts to be expired and future demands. The literature suggests online heuristics that consider the current demand for VMs and the purchase history in the decision-making about the pricing model. Although these heuristics provide guaranteed minimum performance, the decisions are limited due to the lack of knowledge about future demand, influencing the quality of decisions in a capacity planning context. This work investigates whether integrating time series prediction models with online heuristics can improve capacity planning and reduce resource allocation costs. The results show that online heuristics associated with short- or long-term predictions generated by time series prediction state-of-art models can reduce costs.

Keywords: Cloud Computing, Capacity Planning, Online and Augmented Heuristics.