

GERENCIAMENTO DE CAPACIDADE EM NUVENS COMPUTACIONAIS

Alexandre Maros¹, Rafael Rodrigues Obelheiro².

¹ Acadêmico do Curso de Ciência da Computação – CCT – bolsista PROBIC/UDESC

² Orientador, Departamento de Ciência da Computação – CCT – rafael.obelheiro@udesc.br

Palavras-chave: Computação na Nuvem, Séries Temporais, Previsão de Demanda

Com o surgimento das nuvens computacionais, diversos modelos de serviço foram criados para o cliente utilizar essa tecnologia. Um desses modelos é a infraestrutura como serviço (IaaS), que visa disponibilizar recursos computacionais como CPU, memória e rede para o usuário através de máquinas virtuais e/ou contêineres. No modelo IaaS, o cliente tem controle dos recursos, podendo acomodar o poder computacional às suas necessidades, pagando apenas pelos recursos necessários, assim gerando um menor custo final. No entanto, escolher a melhor combinação de recursos não é uma tarefa fácil já que envolve um conhecimento prévio extenso da aplicação a ser implantada e de como esse serviço do cliente será utilizado.

O que ocorre quando o cliente fica encarregado de escolher esses recursos é ou o subprovisionamento ou o superprovisionamento. O subprovisionamento é a falta dos recursos necessários para a aplicação ser executada sem afetar seu desempenho, já o superprovisionamento é justamente reservar recursos acima do necessário. Em ambos os casos há prejuízos para o cliente: no subprovisionamento a aplicação não desempenhará corretamente a função esperada e no superprovisionamento, embora a aplicação execute corretamente, o cliente acaba pagando mais do que o necessário. O normal é realizar o superprovisionamento para evitar problemas, já que afetar o desempenho de uma aplicação é geralmente pior.

Com base nisso, a presente pesquisa tem como principal objetivo a análise do uso de séries temporais para a previsão de demanda de recursos de rede em nuvens computacionais, avaliando sua utilidade nesse cenário. Para isso, inicialmente realizou-se uma revisão bibliográfica dos principais métodos e modelos utilizados para a previsão de séries temporais (Suavização Exponencial Simples, Holt, Holt-Winters, ETS e ARIMA) junto com os trabalhos relacionados da área, em especial, o trabalho de (MOREIRA JR; OBELHEIRO, 2016), onde a mesma técnica de previsão é aplicada para recursos de CPU e memória.

Com base na revisão bibliográfica efetuada, a etapa de modelagem dos dados foi efetuada. Para isso, utilizou-se o *software* estatístico R junto com o pacote *forecast*. Os dados referentes ao uso da rede (taxa de recepção e taxa de transmissão) pertencem a dois servidores em produção, aqui denominados como ZMU3 e ZMU4. Inicialmente, os dados foram analisados e um período de amostragem de 22 dias foi selecionado, sendo que 15 dias foram utilizados como uma série de treinamento, onde as observações são utilizadas para ajustar os modelos, e os últimos 7 dias foram utilizados como uma série de teste, que visa testar a efetividade do modelo ajustado.

Para verificar o quão bem os modelos se ajustaram aos dados, três métricas propostas por (MOREIRA JR; OBELHEIRO, 2016) foram utilizadas: Subestimativa média (SM), percentual de observações não agregadas acima da previsão (POAP) e a previsão acumulada relativa (PAR). As duas primeiras métricas são utilizadas para quantificar as subestimativas de um modelo (uma subestimativa geralmente é pior que uma superestimativa) e a terceira tenta quantificar o custo de cada modelo. Para exemplificar, a Figura 1 mostra as métricas POAP e SM para as previsões referentes ao servidor ZMU4 e a Figura 2 mostra como as previsões se comportaram com base na série de testes.

método	rx		tx	
	POAP	SM	POAP	SM
SES	2.93	420.84	3.68	1688.04
Holt	3.52	378.28	3.57	1718.75
Holt-Winters	2.13	342.69	4.46	1345.82
ETS	3.52	349.65	4.33	1472.90
ARIMA	3.57	387.00	3.53	1717.64

Fig. 1 POAP (%) e SM (kB/s) para o servidor ZMU4

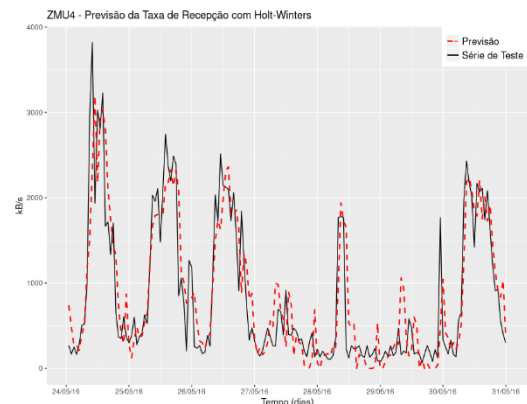


Fig. 2 Previsão da Taxa de Recepção com o método Holt-Winters

Em geral, os métodos e modelos conseguiram gerar uma previsão satisfatória para o uso de banda para ambos os servidores, sendo que em média apenas 3.52% das observações ficaram acima da previsão para o servidor ZMU4 (métrica POAP da Figura 1), ou seja, a previsão gerou poucas subestimativas. Os trabalhos futuros envolvem analisar com mais detalhes os modelos gerados, verificando possíveis melhorias que podem ser realizadas visando capturar mais características das séries e consequentemente gerar melhores previsões. Também há a possibilidade de buscar e aplicar diferentes modelos de previsão que não foram inicialmente abordados.

Referências

MOREIRA JR, D. A.; OBELHEIRO, R. R. Previsão de demanda de recursos em nuvens IaaS usando séries temporais. In: **Simpósio Brasileiro de Engenharia de Sistemas Computacionais (SBESC)**. João Pessoa, PB: [s.n.], 2016