

Mineração de bancos de dados não relacionais usando inteligência artificial para análise de log de nuvem privada de contêineres

RESUMO EXPANDIDO - Disciplina de TCC290009

Maria Luiza Theisges

Estudante do Curso de Engenharia de Telecomunicações

Ederson Torresini

Professor orientador

Semestre 2018-1

Resumo- *A computação em nuvem propiciou que os aplicativos fossem utilizados em nuvem por meio de uma conexão de internet e trouxe o aumento da escalabilidade para diversos tipos de demanda. A Coordenadoria de Tecnologia da Informação e Comunicação (CTIC) do Instituto Federal de Santa Catarina (IFSC) câmpus São José vem trabalhando de forma a disponibilizar todas as aplicações oferecidas pelo campus em nuvem de contêineres, visando diminuir a redundância destes serviços e aumentar a eficiência e escalabilidade do sistema, na forma de microsserviços. Através do uso de aprendizado de máquina, pretende-se, a partir da avaliação dos logs armazenados no banco de dados, desenvolver uma ferramenta que auxilie na análise e na obtenção de indicadores de eficiência da nuvem do Instituto.*

Palavras-chave: Inteligência artificial. Banco de dados. Contêiner.

1 Introdução

Com o surgimento e a evolução da computação em nuvem, emergiu uma nova arquitetura de *software* que propiciou a criação de aplicações baseadas em microsserviços, permitindo o aumento da escalabilidade para atender aos diversos tipos de demanda existentes (KUSZKA, 2017). Com o crescimento de aplicações executando em nuvem, surgiram tecnologias de fácil desenvolvimento e manutenção como, por exemplo, os contêineres. Os contêineres podem ser considerados um novo tipo de virtualização e são mais versáteis que as máquinas virtuais (MANDIC, 2018).

Um contêiner disponibiliza os recursos básicos necessários para o funcionamento de uma aplicação, sendo esta executada em um ambiente isolado. Com a abstração do *hardware* e do sistema operacional, junto com o auxílio do microsserviço, cria-se a portabilidade do contêiner a qual possibilita a execução deste em uma máquina física ou virtual, e em nuvem privada ou pública, além de prover o desenvolvimento e infraestrutura mais ágeis (KUSZKA, 2017). Isto é possível uma vez que, em contêineres, as aplicações geralmente não fazem uso de *hardware* específico, podendo ser facilmente portadas.

O Instituto Federal de Santa Catarina - câmpus São José (IFSC/SJ) é um exemplo que atende a diversas pessoas diariamente, entre professores, alunos e funcionários e oferece a cada um destes públicos aplicações locais, como *Matlab* e *Quartus*, e também aplicações *Web*, como blog e *Wiki* (REIS, 2017; FILHO, 2012). Porém, existem serviços redundantes, ocasionando pouca eficiência no uso dos recursos e má alocação dos mesmos, que são divididos em diversas máquinas virtuais nas quais nenhum recurso é compartilhado. Com isso, acontecem gargalos de recurso na rede. Reis (2017, p. 31) afirma que "Para que esses serviços funcionem em nuvem com qualidade de serviço, é preciso analisar a infraestrutura do IFSC em seus três componentes básicos: rede, processamento e armazenamento de dados."

A CTIC está trabalhando na possibilidade de executar todos os serviços disponibilizados pelo IFSC/SJ em contêineres, melhorando a distribuição dos recursos. Com isto, pretende-se facilitar a manutenção e o custo que se tem com licenças de *software*, por exemplo. Porém, atualmente, não há recursos que possibilitem a CTIC verificar a eficiência da infraestrutura existente no câmpus. Deste modo, com este trabalho pretende-se criar uma ferramenta ou um método que auxilie na análise da eficiência dos recursos presentes na infraestrutura do câmpus.

2 Metodologia

O trabalho se inicia com estudos sobre a infraestrutura da CTIC, como forma de conhecer o cenário atual do câmpus. Posteriormente, serão analisadas as técnicas de inteligência artificial para entender sobre o funcionamento de cada uma. O passo seguinte será selecionar aquelas que vão ser utilizadas no trabalho na parte de mineração de dados. No desenvolvimento do trabalho, serão analisados os *logs* da nuvem do IFSC/SJ como forma de criar indicadores de eficiência de infraestrutura e, com isto, averiguar se atualmente ela está sendo eficiente. As técnicas, previamente selecionadas, serão aplicadas em conjunto com os *logs* armazenados em um banco de dados não relacional. Os resultados obtidos a partir da mineração destes dados, serão avaliados com a equipe da CTIC para determinar se são, de fato, compatíveis com a realidade da infraestrutura atual do Instituto.

3 Resultados e Discussão

Como o volume de dados gerado pela nuvem de contêineres do câmpus é gigantesco, é difícil para a equipe da CTIC determinar o quão eficiente é a infraestrutura atual do IFSC/SJ, uma vez que estes dados devem ser primeiro analisados para que seja possível afirmar tal capacidade. O que se tem hoje, devido a vivência da equipe, é conseguir indicar a eficiência com base na experiência, mas sem muita certeza. Para mensurar de forma mais exata quão eficiente é a infraestrutura atualmente, serão analisadas técnicas de inteligência artificial para sintetizar indicadores de eficiência da infraestrutura do câmpus, possibilitando que isto seja usado de forma contínua pela CTIC.

4 Considerações Parciais/Finais

Através do aprendizado de máquina é possível obter as tendências da base de dados e realizar uma síntese das informações de forma a gerar os indicadores de eficiência. Pretende-se, com a base de dados existente, fazer uma análise temporal juntamente com a correlação de valores para obter os indicadores e, com isto, conseguir encontrar padrões existentes e poder discriminar os valores ressonantes no sistema analisado.

Referências

FILHO, A. A. G. M. *Infraestrutura de redes de computadores focada em SaaS*. [S.l.], 2012. Disponível em: <https://wiki.sj.ifsc.edu.br/wiki/images/4/4b/TCC_AlexandreAugusto.pdf>. Acesso em: 18 de abril de 2018.

KUSZKA, B. *Desmistificando os containers*. 2017. Disponível em: <<https://canaltech.com.br/infra/desmistificando-os-containers/>>. Acesso em: 25 de abril de 2018.

MANDIC. *O que é IaaS, SaaS, tipos de Cloud e Virtualização*. 2018. Disponível em: <<https://www.mandic.com.br/artigos/o-que-e-iaas-tipos-cloud-virtualizacao/>>. Acesso em: 25 de abril de 2018.

REIS, R. B. *Teste de desempenho de sistemas de arquivos distribuídos em nuvem computacional privada no IFSC câmpus São José*. [S.l.], 2017. Disponível em: <https://wiki.sj.ifsc.edu.br/wiki/images/c/cd/TCC_Raphael_Brito_Reis.pdf>. Acesso em: 18 de abril de 2018.