

Técnicas para Gestão e Análise de Bancos de Dados em Nuvem

Leonardo O. Moreira¹, Carlos S. S. Marinho¹, José S. Costa Filho¹,
Flávio R. C. Sousa¹, Javam C. Machado¹

¹Laboratório de Sistemas e Banco de Dados (LSBD)
Universidade Federal do Ceará (UFC) – Fortaleza, CE – Brasil

{leonardo.moreira, sergio.marinho, serafim.costa}@lsbd.ufc.br,

{flavio.sousa, javam.machado}@lsbd.ufc.br

Abstract. *Many applications are data-driven and manage large volumes of information. In this sense, Database Management Systems (DBMSs) are potential candidates for services in computational clouds. Several works are being proposed as techniques to improve the databases' quality of services implanted in the computational clouds. This paper discusses some works that demonstrate techniques for management and data analysis involving relational databases implanted in the cloud.*

Resumo. *Muitas aplicações são orientadas a dados e fazem a gestão de grandes volumes de informação. Neste sentido, Sistemas Gerenciadores de Bancos de Dados (SGBDs) são potenciais candidatos para serem serviços em nuvens computacionais. Diversos trabalhos estão sendo propostos como técnicas para melhorar a qualidade de serviços dos bancos de dados implantados nas nuvens computacionais. Este artigo comenta sobre alguns trabalhos que demonstram técnicas para gestão e análise de dados envolvendo bancos de dados relacional implantados em nuvem.*

1. Introdução e Motivação

Computação em Nuvem é uma solução tecnológica que fornece recursos computacionais, sob demanda, com qualidade de serviço (QoS) e pagamento baseado no uso [Sousa et al. 2018]. A qualidade de serviço está relacionada ao conceito de *Service Level Agreement* (SLA), onde o usuário e provedor estabelecem um contrato de qualidade no uso das nuvens. Por exemplo, o usuário pode estabelecer que o tempo de resposta de requisições a sua aplicação não deve ser superior a um limite estipulado.

Como muitas aplicações são orientadas a dados, Sistemas Gerenciadores de Bancos de Dados (SGBDs) são potenciais candidatos para implantação nas nuvens computacionais [Marinho et al. 2017]. Diante disso, alguns trabalhos [Moreira et al. 2014] [Marinho et al. 2017] [Sousa et al. 2018] estão surgindo com técnicas e soluções para melhorar a QoS de bancos de dados implantados em nuvens computacionais. Este artigo tem como objetivo principal comentar sobre algumas técnicas que são utilizadas para gestão e análise de dados nos bancos de dados em nuvem.

2. Técnicas de Gestão e Análise de Dados para Bancos de Dados em Nuvem

Nesta seção, são comentados alguns trabalhos que fazem gestão dos serviços de bancos de dados em nuvens computacionais para melhorar a QoS em relação a redução de violações

de SLA. Moreira et al. (2014) propuseram uma solução de migração de bancos de dados relacional em nuvem. A solução utilizou técnicas de predição de cargas de trabalho de bancos de dados para estimar o instante de tempo de possíveis violações de SLA. Para isso, adotou-se dois métodos de predição: *AutoRegressive Integrated Moving Average* (ARIMA) e *Support Vector Regression* (SVR). Por meio dos experimentos, é possível observar que o ARIMA obteve bons resultados em predições de curta observação, já o SVR obteve bons resultados em predições de longa observação.

Marinho et al. (2017) projetaram um serviço de balanceamento de carga preditivo para bancos de dados replicados em nuvem. Utilizou-se o modelo ARIMA para realizar análise e predição das cargas de trabalho dos bancos de dados. O trabalho consiste em receber uma transação e redirecionar para a réplica que está menos sobrecarregada. A réplica considerada menos sobrecarregada é aquela que obtiver menor tempo de resposta, no futuro, por meio das predições com o ARIMA. De acordo com os resultados, o ARIMA obteve bons resultados para predições de cargas de trabalho em janelas de curta observação. Sousa et al. (2018) demonstraram uma abordagem preditiva e elástica para replicação de bancos de dados relacional em nuvem. Utilizou-se o ARIMA para prever os comportamentos das cargas de trabalho dos bancos de dados e observar futuras violações de SLA. Ao observar que uma violação de SLA poderá ocorrer, a abordagem faz provisionamento de réplicas para balancear a carga de trabalho e remoção de réplicas quando houver ociosidade. Experimentos orientados a desempenho foram conduzidos e a abordagem proposta conseguiu reduzir, por meio da predição e replicação, as violações de SLA.

3. Conclusão

Este trabalho apresentou algumas técnicas relacionadas a gestão e análise de dados para bancos de dados em nuvens computacionais. Os trabalhos apresentados utilizam modelos de predição para análise e inferência do comportamento das cargas de trabalho de bancos de dados no intuito de visualizar possíveis violações de SLA e realizar ajustes na infraestrutura a fim de evitar as ocorrências de tais violações. Como técnicas de gestão da infraestrutura foram apresentados trabalhos que utilizam balanceamento de carga, migração e replicação de dados. De acordo com os resultados apresentados em tais trabalhos, os modelos de predição ajudaram a visualizar, de forma antecipada, possíveis violações de SLA.

Referências

- Marinho, C. S. S., Coutinho, E. F., Costa Filho, J. S., Moreira, L. O., Sousa, F. R. C., and Machado, J. C. (2017). A Predictive Load Balancing Service for Cloud-Replicated Databases. In *Proceedings of the 32rd Brazilian Symposium on Databases (SBBD)*, pages 210–215, Uberlândia, Minas Gerais, Brazil. Brazilian Computer Society (SBC).
- Moreira, L. O., Farias, V. A. E., Sousa, F. R. C., Santos, G. A. C., Maia, J. G. R., and Machado, J. C. (2014). Towards improvements on the quality of service for multi-tenant rdbms in the cloud. In *IEEE 30th International Conference on Data Engineering Workshops*, pages 162–169.
- Sousa, F. R. C., Moreira, L. O., Costa Filho, J. S., and Machado, J. C. (2018). Predictive elastic replication for multi-tenant databases in the cloud. *Concurrency and Computation: Practice and Experience*, 0(0):e4437.