

Uma Arquitetura Aberta de Baixo Custo e Demanda Energética para Big Data em um Ambiente de Fog Computing

Bruno Machado Agostinho, Francieli Zanon Boito, Mario Antônio Ribeiro Dantas

Departamento de Informática e Estatística
PPGCC - Universidade Federal de Santa Catarina (UFSC)
Florianópolis – SC – Brasil

{bruno.agostinho, francieli.boito}@posgrad.ufsc.br,
mario.dantas@ufsc.br

***Resumo.** Esta proposta aborda uma pesquisa relativa ao paradigma de Fog Computing, como uma alternativa às arquiteturas tradicionais. Visa também a utilização de um conjunto com dispositivos de baixo custo e demanda energética a fim de executar aplicações Big Data.*

1. Introdução

No contexto da computação distribuída, está se tornando cada vez mais comum ver instituições migrando sua infraestrutura para a nuvem a fim de cortar gastos. Muitas aplicações migradas conseguem manter ou melhorar seu desempenho, no entanto algumas acabam por ficarem prejudicadas ou mais expostas no que se refere à segurança e privacidade de informações. Embora a computação em nuvem seja vista como uma plataforma de serviços promissora para internet, a segurança e privacidade dos dados hoje são consideradas uns dos maiores desafios [Wei et al 2013].

Nos últimos anos, vem surgindo um novo conceito relacionado à computação em nuvem chamado de “Fog Computing”, que pode ser descrito como uma arquitetura complementar à tradicional de computação em nuvem, que fornece serviços parecidos, mas que fica situada entre a nuvem e os dispositivos finais dos usuários [Stojmenovic e Wen 2014]. Segundo os mesmos autores, esse tipo de arquitetura serviria para melhorar o desempenho de aplicações que hoje em dia estejam localizadas totalmente na nuvem, mas que precisam de requisições com latência muito pequena ou de uma distribuição geográfica nos servidores. Esse conceito tem diversas possibilidades de aplicações, uma delas seria utilizando para aplicações IoT, inserida na parte de Little Data (ou Big Stream) [Dastjerdi et al 2016].

Muito tem se estudado sobre formas de economizar energia em data centers. Segundo uma pesquisa feita em 188 Data Centers, pelo Green Grid consortium, uma média de 10% dos servidores não são utilizados, podendo ser desligados ou colocados em modo de baixo consumo a fim de economizar energia enquanto estiverem ociosos [Orgerie, Assuncao e Lefreve 2014]. Uma maneira possível para a redução do consumo seria a utilização de dispositivos de baixa demanda energética substituindo os servidores tradicionais. Esse tipo de computador, muitas vezes chamado de SOC (System on Chip), vem se popularizando e incrementando sua capacidade de processamento. Um exemplo disso é a linha Cortex-A, que teve pelo menos 6 modelos novos lançados entre 2010 e 2015 [Halpern, Zhu e Reddi 2016].

2. Proposta

Segundo [Loghin et al 2015], embora o dimensionamento de desempenho requiera computadores com bons recursos de CPU e I/O, os servidores utilizando processadores ARM se tornam atrativos devido ao seu baixo custo, baixo consumo de energia e tamanho reduzido. O mesmo ainda frisa os benefícios e uma mudança de paradigma na área de Big Data caso seja possível utilizar esses dispositivos.

Atualmente, é possível instalar aplicações que normalmente são utilizadas em servidores tradicionais, como por exemplo o sistema de arquivos distribuído do Hadoop (HDFS), em conjunto com outras aplicações, como Hive e Spark, nestes dispositivos. É nesse contexto que esta pesquisa tem como objetivo promover a junção deles para a execução de aplicações Big Data abertas em uma arquitetura de Fog Computing. Para tal, serão utilizados dispositivos em conjunto para testar a instalação do Hadoop e Spark, verificar a capacidade individual de cada um dos dispositivos e como funcionarão em cluster. Também será feita uma medição do consumo energético da utilização do conjunto, utilizando um aparelho de medição que ficará conectado na entrada das fontes dos SOCs. Ao final será feita uma avaliação do comportamento dos dispositivos na execução de tais aplicações, em comparação ao uso de arquiteturas convencionais, tendo como base informações coletadas no Grid5000.

A proposta consiste em como os dispositivos podem ser utilizados em uma camada diferente da nuvem para serviços que não necessitam de tanto recurso de processamento a fim de, no contexto de de Fog Computing, prover uma solução que seja escalável, segura e de baixa demanda energética, utilizando apenas ferramentas abertas.

Referências

- Wei, L., Zhu, H., Cao, Z., Dong, X., Jia, W., Chen, Y. e Vasilakos, A. V. (2014) Security and privacy for storage and computation in cloud computing. *Information Sciences*, 258, pp 371–386.
- Stojmenovic, I. e Wen, S. (2014) The fog computing paradigm: Scenarios and security issues. *Computer Science and Information Systems (FedCSIS), Federated Conference*.
- Dastjerdi, A. V., Gupta, H., Calheiros, R. N., Ghosh, S. K. e Buyya, R. (2016) Fog Computing: Principles, Architectures, and Applications. *CoRR abs/1601.02752*.
- Orgerie, A.-C., Assuncao, M. D. e Lefreve, L. (2014) A Survey on Techniques for Improving the Energy Efficiency of Large-scale Distributed Systems. *ACM Computing Surveys (CSUR)* 46, pp. 47:1–47:31.
- Halpern, M., Zhu, Y. e Reddi, J. (2016) Mobile CPU's Rise to Power: Quantifying the Impact of Generational Mobile CPU Design Trends on Performance, Energy, and User Satisfaction. *IEEE International Symposium on High Performance Computer Architecture (HPCA)*.
- Loghin, D., Tudor, B. M., Zhang, H., Ooi, B. C. e Teo, Y. M. (2015) A Performance Study of Big Data on Small Nodes. *Proceedings of the VLDB Endowment*, Vol. 8, No. 7.