

Serverless Computing – O data center agora é o computador

Valter Canhizares Filho¹, Renata Spolon Lobato¹

¹Departamento de Ciência de Computação e Estatística – Instituto de Biociências,
Letras Exatas (UNESP) – São Jose do Rio Preto – São Paulo – Brasil

{valter.canhizares, renata.spolon}@unesp.br

Abstract. *Recently there is an increased interest in cloud computing, also leveraged by new serverless computing paradigms. Technology offers greater ease and agility to the developer of any application, in addition to providing cost savings. Thus, the main objective of the study is to investigate the state of art on the subject, list applications and challenges on serverless computing, which raises several discussions and advances in academia and consumer market.*

Resumo. *Recentemente há um aumento no interesse acerca da computação em nuvem, alavancado também pelos novos paradigmas criados pela computação sem servidor. A tecnologia oferece maior facilidade e agilidade ao desenvolvedor de qualquer aplicação, bem como proporcionar redução de custos. Assim, o objetivo principal do estudo é investigar o estado da arte sobre o tema, elencar aplicações e desafios sobre serverless computing, que levanta diversas discussões e avanços na academia e no mercado consumidor.*

1. Introdução

Nos últimos anos, transcorreram-se diversos avanços no campo de computação distribuída, computação em nuvem e sem servidor. O mercado de computação em nuvem espera um crescimento de mais de 445 bilhões de dólares em 2021 para aproximadamente 950 bilhões de dólares até 2026, tendo um crescimento composto anual de mais de 16% no período (MarketsandMarkets, 2021).

A computação em nuvem tem sido objetivo de diversas pesquisas desde o começo do século por oferecer as empresas e desenvolvedores servidores remotos hospedados na internet para processar e gerenciar dados. Com o aumento do volume de dados e aplicações, tornou-se necessário adotar alternativas para reduzir o investimento em bens de capital (CAPEX – Capital Expenditure) bem como as despesas operacionais (OPEX – Operational Expenditure) de uma empresa.

Diante da prerrogativa de economia de recursos computacionais, assim como tempo de desenvolvimento de aplicações e soluções, a computação sem servidor (*serverless computing*) emergiu junto ao crescimento da computação em nuvem, encaminhando a comunidade a novos paradigmas e desafios (Sewak, M., & Singh, S., 2018).

O objetivo da presente pesquisa é demonstrar as características, aplicações e desafios desta modalidade de computação. A seção 2 cita as características, seguido da seção 3 que apresenta cenários de aplicações. A seção 4 trata dos desafios e, por fim, na seção 5 a conclusão e trabalhos futuros.

2. Computação em Nuvem e sem Servidor

A computação em nuvem tem como características ser altamente distribuída, de larga escala e elástica, apresentando ao usuário uma computação ubíqua, transparente e de interação consistente e escalável (Tanenbaum, 2007). Na mesma lógica, provedoras desse tipo de serviço oferecem diferentes modelos para o mercado consumidor e possibilitam assinatura de capacidade computacional sob demanda para receber, processar e armazenar dados.

Os modelos de computação em nuvem são previstos por Mell, P. & Grance, T. (2011) em três pilares, que representam três níveis de abstração diferentes: *Software* como Serviço (SaaS), Plataforma como Serviço (PaaS) e Infraestrutura como Serviço (IaaS). A computação sem servidor se enquadra no modelo de *Software* como Serviço, pois tem a premissa da não necessidade de gerenciamento do servidor, sistema operacional, armazenamento por parte do desenvolvedor, segundo Van Eyk et al. (2018).

Observa-se na computação sem servidor uma abstração em relação a infraestrutura, configuração e operação do sistema. Nesta modalidade de computação o usuário tem a liberdade usar micro serviços na nuvem, como execução de código sob demanda, aplicações baseadas em eventos e quebra de grandes pedaços de código em pequenas instruções (Sewak, M., & Singh, S., 2018), com custo granular inerente ao uso e execução distribuída (Van Eyk et al, 2018).

Existem dois pontos interessantes a se observar na computação sem servidor: o custo e a elasticidade, com a premissa de se pagar apenas o mínimo necessário para executar instruções e aplicações, geralmente em milissegundos de execução e uso da capacidade computacional. Torna-se um modelo atrativo para execução de códigos ocasionalmente usados, oferece escalada do zero de uma aplicação qualquer. Dois desafios neste sentido seriam agendamento e otimização dos recursos computacionais usados no momento da execução, segundo Castro, P. et al. (2019). A elasticidade é outro ponto a se observar, visto que oferece a possibilidade de escalar de zero para n baseado nas instruções. Em contrapartida é um desafio para a provedora no sentido de oferecer e alocar a capacidade computacional requisitada de modo ágil (Moreira & Barreto, 2019).

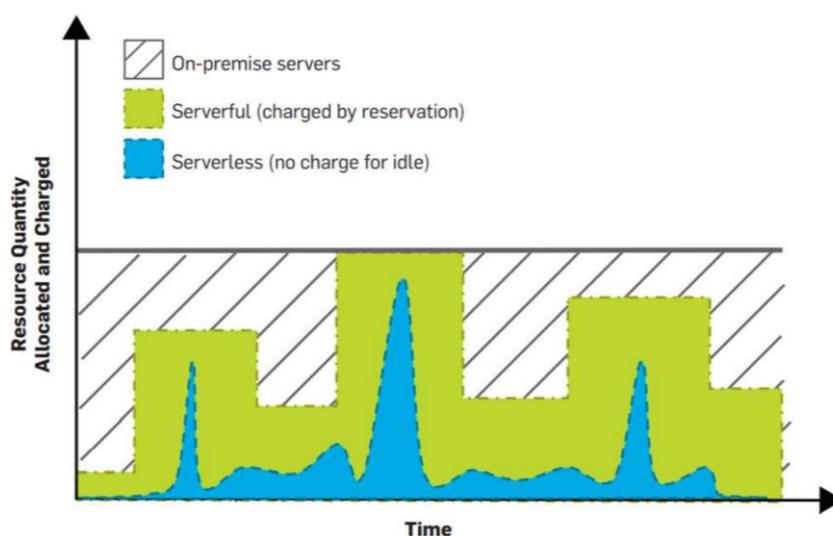


Figura 1 - Adaptação Schleier-Smith et. al (2021). Comparação do custo de capacidade computacional alocada e cobrada por determinada execução.

A Figura 1 ilustra um comparativo entre a alocação de capacidade computacional e seu respectivo custo no que tange a computação em nuvem, computação sem servidor e ambiente *On-premises*, que se trata de um ambiente físico e real que uma empresa tem a responsabilidade de processar as aplicações em seu próprio *hardware* e *software*. É possível notar que a computação sem servidor pode gerar economia dependendo de seu contexto, a modalidade de pagar sob a demanda utilizada é um atrativo.

3. Aplicações com Computação sem Servidor

Segundo relatório de casos de uso da IBM (2021), cerca de 30% da utilização da computação sem servidor é voltada serviços *web* e interfaces de programação de aplicações – APIs, aproximadamente 20% para processamento de dados como extração e transformação de base de dados, 17% para integração de aplicações terceiras e menos de 10% para *Chatbots* e *Internet das Coisas* – IoT. Muitos trabalhos exploram a criação e uso de *frameworks* que atuam junto à computação sem servidor, para tornar o desenvolvimento mais ágil.

Hassan & Barakat (2021) elencam cinco principais benefícios da computação sem servidor como: baixo custo, alta escalabilidade, fácil gerenciamento e facilidade de publicação de aplicações com menor latência. É possível aplicar a computação sem servidor a outras áreas como *chatbots*, a integração com usuários por meio de voz e texto de forma prática, processamento de dados, agindo como um modelo de processamento paralelo ou armazenamento simples com funções de entrada e saída. É possível também agregar a segurança com a criação de sistema de detecção de ameaças com amparo de *Kubernetes*, rede de computadores com adoção de Redes Definidas por *Software* (SDNs) e por fim, *internet das Coisas* (IoT).

4. Desafios da Computação sem Servidor

Tendo em vista os recentes avanços na área de computação sem servidor, fica evidente que novos paradigmas e novos desafios emergem e dão novos horizontes ao uso da tecnologia. Neste caso é possível evidenciar desafios em relação a performance da computação sem servidor, por conta de sobrecargas de chamadas de serviços e agendamentos; desafios no *cold start*, ou arranque à frio da aplicação, no que tange a execução do zero e inicialização dos processos a serem executados; a precificação e o custo, que são desafiadores tanto para provedoras quanto para os desenvolvedores fazendo-se necessário novas métricas no negócio; segurança no isolamento das funções, por ser uma plataforma compartilhada entre vários clientes. Por fim, o *lock-in*, ou aprisionamento forçado da provedora, que é a dependência de contratar um único ecossistema de computação em nuvem para execução das tarefas de desenvolvimento, criando uma amarra com a empresa contratada. Neste caso é preciso pensar em um gerenciamento mais heterogêneo das aplicações.

5. Conclusão e Trabalhos Futuros

É inegável que a computação sem servidor é uma tecnologia promissora e vai abrir espaços para novos paradigmas da computação em nuvem, visto que a agilidade, facilidade e elasticidade são bons pontos fortes dessa tecnologia.

Trabalhos futuros podem emergir entendido que novas tecnologias devem mitigar os desafios que foram pontuados no presente artigo, por meio de *frameworks*, por exemplo,

que agilizam alguns processos. No presente é necessário avançar em relação aos desafios de *lock-in* da provedora bem como o *cold start* que afeta as aplicações. Opções *open source* são bem-vindas para este universo para que os paradigmas da computação sem servidor avancem para atender as demandas do mercado, com premissas de otimização de custo e recursos.

O presente trabalho contribui para melhor entendimento acerca da computação sem servidor e melhor entendimento das possibilidades e desafios desta área. É esperado que essa contribuição ofereça alicerce para novas propostas para os sistemas distribuídos, computação em nuvem e computação sem servidor.

Referências

- Castro, P., Ishakian, V., Muthusamy, V., & Slominski, A. (2019). The rise of serverless computing. *Communications of the ACM*, 62(12), 44–54. doi:10.1145/3368454
- Hassan, H.B., Barakat, S.A. & Sarhan, Q.I. Survey on serverless computing. *J Cloud Comp* 10, 39 (2021). <https://doi.org/10.1186/s13677-021-00253-7>
- IBM. Serverless in the enterprise, 2021. Disponível em: <ibm.com/downloads/cas/ZJLWQOAQ>
- MarketsandMarkets. Cloud Computing Market Size, Share and Global Market Forecast to 2026 | COVID-19 Impact Analysis | MarketsandMarkets. Disponível em: <<https://www.marketsandmarkets.com/Market-Reports/cloud-computing-market-234.html>>. Acesso em: 05 nov. 2021.
- Mell, P., & Grance, T. (2011). SP 800-145. The NIST Definition of Cloud Computing. Doi: 10.6028/NIST.SP.800-145
- Moreira, R. D., & Barreto, P. S. (2019). A Framework for Improving Cold Start Time in Function-as-a-service (FaaS). 2019 XLV Latin American Computing Conference (CLEI). doi:10.1109/clei47609.2019.235112
- Schleier-Smith, J., Sreekanti, V., Khandelwal, A., Carreira, J., Yadwadkar, N. J., Popa, R. A., ... Patterson, D. A. (2021). What serverless computing is and should become. *Communications of the ACM*, 64(5), 76–84. doi:10.1145/3406011
- Sewak, M., & Singh, S. (2018). Winning in the Era of Serverless Computing and Function as a Service. 2018 3rd International Conference for Convergence in Technology (I2CT). doi:10.1109/i2ct.2018.8529465.
- Tanenbaum, A. S. *Sistemas Distribuídos: Princípios e Paradigmas*. 2ª edição, Pearson – Prentice Hall, 2007. ISBN: 0-13-239227-5.
- Van Eyk, E. Toader, L., Talluri, S., Versluis, L., Uta, A., & Iosup, A., Serverless is More: From PaaS to Present Cloud Computing. *IEEE Internet Computing*, 22(5), 8–17. doi:10.1109/mic.2018.053681358