

# Computação em Nuvem Autônoma: Oportunidades e Desafios

Flávio R. C. Sousa<sup>1</sup>, Leonardo O. Moreira<sup>1</sup>, Javam C. Machado<sup>1</sup>

<sup>1</sup> Mestrado e Doutorado em Ciência da Computação (MDCC)  
Universidade Federal do Ceará (UFC) – Fortaleza, CE – Brasil

{sousa, leoomoreira, javam}@ufc.br

**Abstract.** *Cloud computing is a recent trend of technology aimed at providing on-demand services with payment based on usage. However, this computing model requires major technological changes, especially in the autonomous management because the cloud environment presents a great number and variety of resources. Service providers, in turn, need to address issues of scalability, availability, performance, and cost. This paper aims to discuss cloud computing and autonomic management, highlighting opportunities and research challenges in this context.*

**Resumo.** *Computação em nuvem é uma tendência recente de tecnologia cujo objetivo é proporcionar serviços sob demanda com pagamento baseado no uso. Entretanto, este modelo de computação requer grandes mudanças tecnológicas, principalmente no gerenciamento autônomo, pois o ambiente em nuvem apresenta uma grande quantidade e variedade de recursos. Os provedores de serviço, por sua vez, necessitam tratar aspectos de escalabilidade, disponibilidade, desempenho e custo. Este artigo tem como objetivo discutir sobre a computação em nuvem e o gerenciamento autônomo, destacando oportunidades e desafios de pesquisa neste contexto.*

## 1. Introdução

Computação em nuvem fornece infraestrutura, plataforma e software como um serviço sob demanda com pagamento baseado no uso [Buyya et al. 2009]. Com isso, as empresas e os desenvolvedores não precisam realizar grandes investimentos em hardware e manutenção para implementar serviços, permitindo mais foco em inovação e na melhoria dos negócios das empresas [Sousa et al. 2010]. Por outro lado, os ambientes de nuvem são inerentemente grandes, complexos, heterogêneos e altamente dinâmicos e os provedores devem tratar questões de qualidade do serviço, disponibilidade e eficiência energética. Os níveis de complexidade na nuvem excedem a capacidade humana, forçando a retirada das pessoas do processo de gerência e aumentando a autonomia dos sistemas. Além disso, soluções para o gerenciamento neste contexto devem se adaptar as mudanças e comportamentos da nuvem de acordo com instruções especificadas em alto nível pelos administradores destas soluções.

A computação autônoma é inspirada em sistemas biológicos para lidar com desafios de complexidade, dinamismo e heterogeneidade [Kephart and Chess 2003], características presentes nos ambientes de computação em nuvem e, assim, fornece uma abordagem promissora neste contexto. Embora a computação em nuvem apresente certas características autônomas como o provisionamento automático de recursos, seu objetivo é reduzir o custo dos recursos ao invés de reduzir a complexidade do sistema

[Zhang et al. 2010]. Na próxima seção apresentamos oportunidades e desafios de pesquisa no gerenciamento autônomo do ambiente de computação em nuvem e, na última seção, as conclusões deste trabalho.

## **2. Oportunidades e Desafios**

A computação em nuvem é um paradigma cada vez mais popular e tem atraído esforços da indústria e da academia. Associado a isto, a computação autônoma permite melhorar os esforços já aplicados neste contexto [CometCloud 2011]. A seguir destacamos alguns desafios e oportunidades de pesquisa em computação em nuvem autônoma.

### **2.1. Desenvolvimento de Sistemas Autônomos para Nuvem**

A nuvem é um sistema autônomo onde hardware e software podem ser automaticamente reconfigurados, orquestrados e estas modificações são apresentadas de forma transparente para os usuários. Neste contexto, diferentes modelos evoluíram rapidamente para aproveitar as tecnologias de software, plataformas de programação, armazenamento de dados e infraestrutura de hardware como serviços. Enquanto estes modelos se referem ao núcleo dos serviços de computação em nuvem, suas inter-relações têm sido ambíguas e a viabilidade de sua interoperabilidade é questionável. Além disso, cada serviço de computação em nuvem possui *interfaces* e protocolos diferentes [Youseff et al. 2008].

O gerenciamento do ambiente de nuvem possui algumas características que diferem de outros ambientes, dentre as quais podemos destacar a intervenção humana limitada, carga de trabalho altamente variável e uma grande quantidade e variedade de recursos compartilhados. De acordo com as requisições dos usuários, estes recursos são expandidos ou reduzidos para manter a qualidade do serviço. Associado a isso, os provedores de serviços em nuvem devem tratar aspectos de pagamento baseado no uso enquanto utilizam recursos compartilhados para vários usuários. Com isso, tem-se um ambiente dinâmico, elástico e distribuído, o que dificulta o desenvolvimento de soluções para o gerenciamento destes ambientes.

Um sistema autônomo para nuvem deve monitorar o comportamento e desempenho do ambiente, tratar questões de tolerância a falhas, elasticidade e balanceamento da carga de trabalho, modelar e prever o comportamento para as cargas de trabalho e realizar ações para lidar com as variações destas cargas. Técnicas de aprendizagem de máquina podem ser utilizadas para classificar a carga de trabalho e prever o custo de operações, melhorando o provisionamento [Agrawal et al. 2011]. Questões de segurança, privacidade e confidencialidade também devem ser abordadas, visto que os elementos autônomos colaboram para realizar suas atividades. Por fim, é importante definir uma forma de integrar estes novos sistemas autônomos aos sistemas existentes.

### **2.2. Políticas para Computação em Nuvem**

As políticas devem representar as propriedades dos sistemas autônomos e o comportamento, objetivos, adaptação e interação entre os elementos autônomos, assim com a influência externa e a demanda do sistema. Assim, as políticas devem ser simples, permitindo a intervenção humana de forma precisa e evitando a definição de objetivos inconsistentes. Para tanto, pode-se desenvolver modelos matemáticos e algoritmos que, de acordo com objetivos de alto nível, auxiliem a garantir os objetivos [Paton et al. 2009]. Como a

quantidade de recursos gerenciados pelos agentes na nuvem é muito grande, um desafio é garantir que o comportamento resultante dos objetivos individuais de cada agente irá conduzir à realização de um objetivo comum.

No ambiente de nuvem, as políticas, regras ou requisitos são orientados ao negócio. Contudo, mas não existe uma padronização destas regras. O grau de automação, abstração para o usuário e customização das políticas e regras que regem um serviço pode variar bastante. Alguns sistemas oferecerem aos usuários a possibilidade de construção de condições simples com base em certas métricas, por exemplo, CPU, memória, enquanto outros utilizam métricas no nível do serviço, por exemplo, relação custo/benefício e permitem estratégias mais complexas [Vaquero et al. 2011]. Dessa forma, um desafio é conceber políticas de alto nível que capturem tanto aspectos técnicos como os objetivos do negócio. As abordagens para o desenvolvimento de políticas baseadas em ontologias podem ser uma alternativa [Strassner et al. 2009]. Ontologias podem ser utilizadas para a organização do domínio de conhecimento de computação em nuvem, seus componentes e suas relações, ajudando a comunidade na melhor compreensão das tecnologias.

### 2.3. Monitoramento de Ambientes em Nuvem

O monitoramento é essencial para o ambiente de computação em nuvem, visto que este ambiente está se alterando constantemente de forma a atingir os objetivos. Para garantir a qualidade do serviço, utiliza-se a abordagem baseada em acordo de nível de serviço (SLA), que contem informações sobre os níveis de disponibilidade, desempenho e penalidades em caso de violação destes níveis. No ambiente em nuvem, o objetivo é minimizar a quantidade de recursos necessários para garantir a qualidade do serviço, o que reduz os custos. Dessa forma, uma questão é como gerenciar de forma automática os recursos disponíveis e a carga de trabalho do sistema para garantir a qualidade do serviço e melhorar a utilização destes recursos. Técnicas de monitoramento adaptativas e dinâmicas deverão ser desenvolvidas para tratar esta questão.

De acordo com o modelo de serviço (IaaS, PaaS ou SaaS) utilizado, o gerenciamento autônomo é diferente. Por exemplo, um provedor IaaS, em geral, gerencia apenas recursos da sua infraestrutura, ao passo que, um provedor de PaaS ou SaaS pode precisar gerenciar a utilização dos recursos em diferentes infraestruturas. Dessa forma, no modelo de PaaS e, principalmente, no modelo de SaaS deve-se ter soluções de monitoramento para melhorar a utilização dos recursos e verificar a garantia dos serviços considerando diferentes infraestruturas.

Outro ponto importante é obter informações sobre cada elemento, já que nem todos os dispositivos possuem uma *interface* padronizada. [Lim et al. 2009] discutem sobre a utilização de sensores e atuadores em ambientes de computação em nuvem e apresentam algumas questões de pesquisa, tais como obter informações sobre os componentes do ambiente. Para monitorar a nuvem, [Lee et al. 2010] propõem uma abordagem para o monitoramento deste ambiente baseada em redes de sensores. Técnicas como *MapReduce* podem ser utilizadas para o processamento do grande volume de dados coletados pelos sensores, bem como para construir uma base de conhecimento.

## 3. Conclusão

Este artigo apresentou problemas e oportunidades para o gerenciamento autônomo de ambientes de computação em nuvem. O desenvolvimento de sistemas autônomos e as

características do monitoramento foram discutidos e alguns desafios foram detalhados. Este artigo não tem a pretensão de descrever exaustivamente as técnicas existente para a construção de ambientes de nuvens com autonomia de gerenciamento. Outras técnicas são igualmente viáveis. O grupo de *Cloud Computing* da UFC participa intensamente na especificação e na construção de um sistema autônomo para o ambiente de nuvem.

## Referências

- Agrawal, D., Abbadi, A. E., Das, S., and Elmore, A. J. (2011). Database scalability, elasticity, and autonomy in the cloud - (extended abstract). In *Database Systems for Advanced Applications - 16th International Conference, DASFAA 2011*, pages 2–15.
- Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., and Brandic, I. (2009). Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Gener. Comput. Syst.*, 25(6):599–616.
- CometCloud (2011). *CometCloud: An Autonomic Cloud Engine*. [www.cometcloud.org](http://www.cometcloud.org).
- Kephart, J. O. and Chess, D. M. (2003). The vision of autonomic computing. *Computer*, 36(1):41–50.
- Lee, K., Murray, D., Hughes, D., and Joosen, W. (2010). Extending sensor networks into the cloud using amazon web services. In *Proceedings of the 1st IEEE International Conference on Networked Embedded Systems for Enterprise Applications, NESEA 2010*, pages 1–7. IEEE.
- Lim, H. C., Babu, S., Chase, J. S., and Parekh, S. S. (2009). Automated control in cloud computing: challenges and opportunities. In *Proceedings of the 1st workshop on Automated control for datacenters and clouds, ACDC '09*, pages 13–18, New York, NY, USA. ACM.
- Paton, N. W., Aragão, M. A. T., Lee, K., Fernandes, A. A. A., and Sakellariou, R. (2009). Optimizing utility in cloud computing through autonomic workload execution. *IEEE Data Eng. Bull.*, 32(1):51–58.
- Sousa, F. R. C., Moreira, L. O., Macêdo, J. A. F., and Machado, J. C. (2010). *Gerenciamento de Dados em Nuvem: Conceitos, Sistemas e Desafios*, pages 101–130. In: PEREIRA, A. C. M.; PAPPA, G. L.; WINCKLER, M.; GOMES, R. L. (Org.). *Tópicos em Sistemas Colaborativos, Interativos, Multimídia, Web e Bancos de Dados, SWIB 2010*, 1. ed. SBC, Belo Horizonte.
- Strassner, J., de Souza, J. N., van der Meer, S., Davy, S., Barrett, K., Raymer, D., and Samudrala, S. (2009). The design of a new policy model to support ontology-driven reasoning for autonomic networking. *J. Network Syst. Manage.*, 17(1-2):5–32.
- Vaquero, L. M., Rodero-Merino, L., and Buyya, R. (2011). Dynamically scaling applications in the cloud. *SIGCOMM Comput. Commun. Rev.*, 41:45–52.
- Youseff, L., Butrico, M., and Da Silva, D. (2008). Toward a unified ontology of cloud computing. In *Grid Computing Environments Workshop, 2008. GCE '08*, pages 1–10.
- Zhang, Q., Cheng, L., and Boutaba, R. (2010). Cloud computing: state-of-the-art and research challenges. *Journal of Internet Services and Applications*, 1:7–18.