



## GERÊNCIA DE CAPACIDADE BASEADA EM MÉTRICAS DE NEGÓCIO PARA APLICAÇÕES DE SOFTWARE COMO UM SERVIÇO

Lília Rodrigues Sampaio<sup>1</sup>, Raquel Vigolvinho Lopes<sup>2</sup>

### RESUMO

A Computação na Nuvem se propõe a fornecer serviços em vários níveis, desde o armazenamento de dados e provimento de recursos computacionais sob-demanda até o acesso a um software na Internet. Em se tratando do serviço de “Software como um serviço” (do inglês, SaaS - Software as a Service), é importante gerenciar a capacidade das aplicações oferecidas. Utilizam-se duas técnicas específicas para gerenciamento a longo e curto prazo dos recursos, o planejamento de capacidade e o escalonamento de aplicações, respectivamente. Este trabalho foi desenvolvido nesse contexto, tendo sido organizado em duas etapas. Primeiro, foram realizadas atividades de verificação e testes em um simulador desenvolvido com a finalidade de apoiar experimentos na área de planejamento de capacidade. Na segunda etapa, iniciamos a avaliação de um novo serviço no mercado dos provedores na nuvem, o escalonamento automático de aplicações. Essa técnica consiste em escalonar dinamicamente a quantidade de recursos, baseado em condições definidas pelo usuário e especificadas considerando métricas coletadas por um monitor ligado à aplicação. Dado que diversos algoritmos de provisionamento dinâmico foram propostos para utilização dentro das técnicas de planejamento e escalonamento, decidimos avaliar a possibilidade de implementação de alguns dos mais importantes neste serviço, bem como suas vantagens e desvantagens considerando métricas de negócio.

**Palavras-chave:** computação na nuvem, escalonamento automático, provisionamento e métricas de negócio

### CAPACITY PLANNING BASED ON BUSINESS METRICS FOR APPLICATIONS OF SOFTWARE AS A SERVICE

### ABSTRACT

Cloud Computing has the purpose of providing services in different levels, making possible since the data storage and computing resources on demand, until the access to a software through the Internet. Regarding Software as a Service (SaaS), it's important to manage the capacity of the applications provided. In order to manage such the resources available it should be used specific techniques in the short and long term, called applications scheduling and capacity planning, respectively. This research was developed in this context, and it was organized in two steps. Firstly, verification activities and tests were performed in a simulator developed with the finality of supporting experiments in the capacity planning area. In the second step, we started an evaluation of a new service identified in the market of cloud providers as applications auto scaling. This technique consists in scheduling dynamically the amount of resources, based on conditions defined by the user itself, and that are specified considering metrics collected by a monitor connected to the application. Since there are a diversity of provisioning algorithms used in combination with the capacity planning and applications scheduling techniques, we decided to evaluate the viability of implementation of the most important ones using this new auto scaling service, as well as the advantages and disadvantages of using this service considering business metrics.

**Keywords:** cloud computing, auto scaling, provisioning and business metrics

---

<sup>1</sup> Aluna do Curso de Ciência da Computação, Centro de Engenharia Elétrica e Informática, Departamento de Sistemas e Computação, UFCG, Campina Grande, PB, E-mail:liliars@isd.ufcg.edu.br

<sup>2</sup> Ciência da Computação, Professora. Doutora, Centro de Engenharia Elétrica e Informática, Departamento de Sistemas e Computação, UFCG, Campina Grande, PB, E-mail:raquel@dsc.ufcg.edu.br