

Aplicações Web Elásticas em Computação em Nuvem

Wesley R. Oliveira¹, Eduardo Silva Vieira¹, Mário Meireles Teixeira^{1,2}

¹Laboratório de Sistemas Avançados da Web – LAWS

²Departamento de Informática
Universidade Federal do Maranhão (UFMA) – São Luís, MA – Brasil

edusvieirap@gmail.com, wesley231096@hotmail.com, mario@deinf.ufma.br

Abstract. *This article describes the main concepts of Cloud Computing exposing features, services and advantages of using it. It addresses a developed elastic web application that was hosted in an open source cloud and with the help of a specific tool could simulate a synthetic load to realize the concept of elasticity in action.*

Resumo. *Este artigo descreve os principais conceitos de Computação em Nuvem expondo recursos, serviços e vantagens de utilizá-la. Aborda uma aplicação web elástica desenvolvida, que foi hospedada numa nuvem de código aberto e com o auxílio de uma ferramenta específica pôde simular uma carga sintética para visualizar o conceito de elasticidade em ação.*

1. Introdução

Em [NIST 2011] e [Herbst 2013], define-se que Computação em Nuvem (em inglês, Cloud Computing) como um modelo para permitir acesso a serviços de modo onipresente, conveniente e sob demanda, constituídos em um conjunto compartilhado de recursos de computação configuráveis (p. ex., redes, servidores, armazenamento, aplicativos e serviços), que podem ser rapidamente provisionados e liberados com o mínimo de esforço de gerenciamento ou interação com o prestador de serviços.

Em [Brebner 2012] e [Lorido-Botran 2014], destaca-se que Computação em Nuvem é um serviço que permite disponibilizar recursos computacionais sob demanda, seja de forma automática, seja de forma manual, independentemente de sua localização física. De fato, Computação em Nuvem permite melhor aproveitar os recursos disponíveis em uma infraestrutura computacional, diminuindo desperdício ou sua não utilização e, conseqüentemente, melhorando a qualidade do serviço fornecido, além de reduzir os custos [Kupferman and Silverman 2016].

Vale frisar que a Cloud Computing é o armazenamento de informações, aplicações e demais serviços na área de TI sob uma plataforma online que pode ser acessada de qualquer lugar, a qualquer momento, sem a necessidade de instalar algum programa. O acesso é feito remotamente e o valor de custo é de acordo com o uso dos serviços.

2. Aplicação Web Elástica

Uma Aplicação Web é qualquer programa de computador que executa uma função específica utilizando o navegador como cliente. A aplicação pode ser tão simples quanto

um quadro de mensagens ou um formulário de contato em um site ou tão complexo quanto um processador de texto ou um aplicativo de jogos para vários jogadores que você baixa para o seu celular [Nations 2018].

Aplicações Web podem ser definidas como sistemas desenvolvidos que executam em web hosts e são acessadas pelos usuários através dos browsers (navegadores). Devido ao fato do acesso ser por um browser, a quantidade de usuários/dispositivos que podem acessá-las é potencialmente ilimitada. Neste tipo de sistema não há a necessidade dos usuários baixarem ou instalarem dados para o dispositivo de origem, onde se processa a parte cliente da aplicação.

Uma Aplicação Web Elástica consiste em um sistema computacional que possui a capacidade de aumentar ou diminuir os recursos alocados aos usuários e/ou aplicações, de maneira automática, o que remete ao conceito de um objeto elástico, que pode esticar ou encurtar-se conforme a força aplicada sobre ele. Na prática, tais recursos podem ser adicionados por meio de cópias e/ou pela expansão de recursos já existentes como processador e memória.

3. Vantagens da Computação em Nuvem

Abaixo encontram-se algumas das características que a nuvem oferece e consequentemente as vantagens oferecidas:

- Pay as You Go (Pague somente pelo que usar) – A cobrança é baseada somente no quanto é usado (horas, minutos, quantidade de recursos), não é cobrado nem a mais, nem a menos.
- Visão Única do Sistema – O controlador da nuvem tem a “impressão” de que esta é apenas um computador normal e à medida que são requisitados novos recursos, estes são inseridos na máquina de maneira manual.
- Autoatendimento sob demanda: Novos recursos são alocados à máquina do usuário sem a necessidade de intervenção humana, apenas pelo reconhecimento automático já pré-configurado pelo usuário, desta maneira todo o hardware e software é reconfigurado e a nova configuração é apresentada ao usuário de maneira transparente.
- Pool de recursos: Os recursos computacionais (físicos ou virtuais) do provedor são divididos em pools para que possam atender a múltiplos usuários simultaneamente. Esses recursos são alocados e realocados dinamicamente, de acordo com a demanda dos usuários.
- Elasticidade: Consiste em que o sistema computacional possui a capacidade de aumentar ou diminuir os recursos alocados aos usuários e/ou aplicações, de maneira automática.

4. Por que utilizar Aplicações Web Elásticas?

Utilizando um exemplo cotidiano, pelo menos alguma vez na vida, um usuário da internet se deparou com uma determinada Aplicação Web que estava “lenta”, não fazia ideia do que estava acontecendo, pensou que talvez o problema fosse com seu computador ou dispositivo. Uma das possíveis causas dessa lentidão do sistema é o acesso de vários usuários de forma simultânea que acaba sobrecarregando o servidor e assim tornando o tempo de resposta maior ou até mesmo ocasionando seu desligamento.

Uma possível solução para o problema seria a utilização de uma nuvem para hospedar a Aplicação e fazer uso da elasticidade, desta maneira quando o servidor estivesse próximo da sobrecarga, a carga de trabalho se dividiria ou novos recursos seriam alocados para evitar a lentidão e até mesmo desligamento do servidor.

Devido às características apresentadas e seu custo-benefício, a Computação em Nuvem e mais especificadamente a elasticidade que ela oferece são um grande aliado para redução de custos, infraestrutura e tempo, podendo ser aplicada nas áreas da indústria, governo, pesquisa etc.

5. Metodologia

Para elaboração do presente artigo foram realizadas pesquisas bibliográficas em livros, artigos e páginas da web relacionadas ao tema.

Para demonstrar os conceitos aprendidos foi desenvolvido uma aplicação web, seguindo rigorosamente todas as etapas na elaboração de um software, com toda a documentação exigida. A partir disso foi desenvolvida a Aplicação Carcontroller que tem como objetivo auxiliar gerentes e vendedores para a venda e aquisição de veículos automobilísticos.

Em seguida foi necessário hospedar a aplicação dentro da nuvem [Eucalyptus], a nuvem foi escolhida por ser gratuita e derivada da famosa AWS (Amazon Web Service), e configurá-la da maneira adequada para suportar a Aplicação e obter o serviço de elasticidade através do balanceador de carga. Por último, para verificar o desempenho da nuvem houve a necessidade de simular vários testes simultâneos para que a nuvem pudesse ser sobrecarregada e utilizasse o serviço de elasticidade.

Com um auxílio de uma ferramenta específica, chamada [JMeter] foi possível definir a quantidade de usuários para os testes assim como tempo, quantidade de iterações, intervalo entre requisições etc.

Após ser submetida à sobrecarga, a nuvem gerou gráficos ilustrando o horário e a quantidade de recursos utilizados, visto isso pôde-se analisar a elasticidade da nuvem em si, o momento em que os recursos se expandiram e se comprimiram.

6. Resultados

Como resultado obtivemos o desenvolvimento da Aplicação Carcontroller, a instalação da nuvem Eucalyptus e os dados colhidos dos testes efetuados. Para demonstrar a Aplicação desenvolvida, abaixo (Figura 1) é possível mostrar a tela de login.

Com o intuito de demonstrar a elasticidade atuando hospedou-se a ferramenta desenvolvia em uma instância, inicialmente e após os testes ela passou a ser hospedada em três (limite máximo de instâncias estabelecidas). Com o auxílio da ferramenta Jmeter foi possível gerar as cargas e desta forma notar o desempenho das instâncias e a divisão da carga. Ambos podem ser notados com a visualização do gráfico de utilização da CPU.

Em testes iniciais foi definido que se durante o período de tempo de um minuto a utilização da CPU atingisse 65% o balanceador de carga entraria em ação e a carga se dividiria, no entanto pelo motivo da aplicação ter mais métodos gets (recupera informações) do que posts (insere informações), o primeiro utiliza menos CPU que o último, não foi

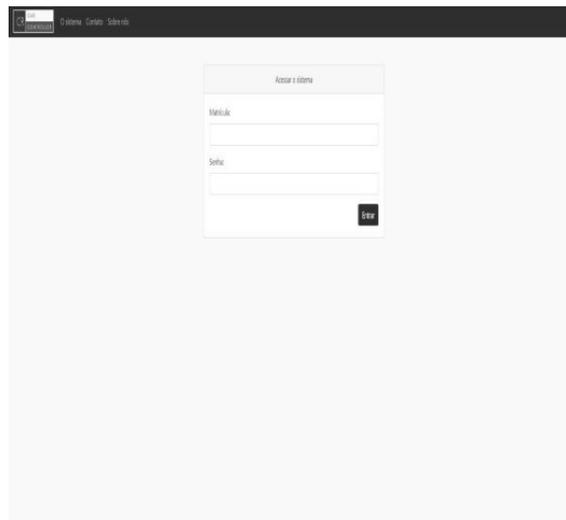


Figure 1. Login do usuário

possível atingir esse limite. A fim de evitar as instâncias operando em modo ocioso foi configurado para serem automaticamente desligadas caso o limite da CPU estabelecido não fosse ultrapassado pelo período de tempo de cinco minutos.

Com base no teste anterior foi mudado o limite de utilização da CPU para 3% e os de tempo permaneceram o mesmo. Destarte, o teste ocorreu da seguinte maneira: a Jmeter gerou cargas que ultrapassaram os limites estabelecidos assim, uma nova instância foi ligada, desse modo a carga foi dividida entre duas instâncias através do balanceador. Ainda com a ferramenta gerando carga, mais uma vez o limite foi atingido e da mesma maneira outra instância foi iniciada, agora com a carga atuando em três instâncias de forma simultânea. Após atingir o limite de instâncias a ferramenta geradora de carga foi desligada e observou-se os gráficos a seguir.

Na figura 2 é possível notar comportamento da CPU por um período de tempo de 3h, este foi o momento em que foi aplicado a carga na aplicação, percebe-se que ao ultrapassar o limite estabelecido durante um minuto a curva começa a diminuir. Com a redução é possível atingir o objetivo de dividir a carga de trabalho entre outras instâncias.

Em seguida, na figura 3 é possível notar a utilização da CPU após o encerramentos dos testes. Por último, na figura 4 e figura 5 é possível observar o momento em que as instâncias são ligadas e desligadas.

7. Conclusão

De acordo com o proposto foi criada uma Aplicação Web elástica e utilizou-se serviços de Computação em Nuvem para hospedá-la, a infraestrutura criada foi sobre uma plataforma de código aberto bem difundida no mundo da Computação em Nuvem.

A aplicação se comportou conforme era esperado e assim pôde usufruir dos serviços oferecidos pela nuvem, principalmente a elasticidade. Os testes foram feitos por meio de geração sintética de carga de usuários através de uma ferramenta específica. As características observadas neste trabalho foram o custo benefício, facilidade, agilidade e robustez apresentados pela nuvem.



Figure 2. Utilização da CPU por um período de 3h

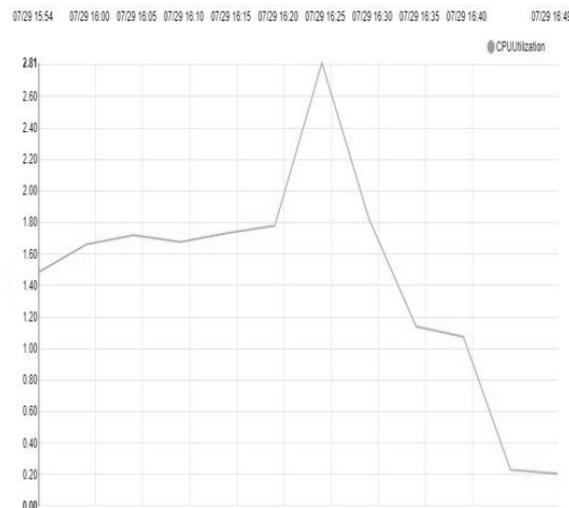


Figure 3. Utilização da CPU após encerramento dos testes

Para trabalhos futuros pretende-se aplicar este estudo em aplicações multimídia, servidores de banco de dados e clusters de servidores.

References

Brebner, P. C. (2012). *Is your cloud elastic enough? Performance modelling the elasticity of infrastructure as a service (iaas) cloud applications*. In Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering, ICPE '12.

Eucalyptus. <https://www.eucalyptus.cloud/>. Online, acessado em: 11/11/2018.

Herbst, N. R. (2013). *Elasticity in cloud computing: What it is, and what it is not*. In Proceedings of the 10th International Conference on Autonomic Computing (ICAC 13).

JMeter. <https://jmeter.apache.org/>. Online, acessado em: 11/11/2018.

Kupferman, J. and Silverman, J. (2016). *Scaling into the cloud. ADVANCED OPERATING SYSTEMS.*

Lorido-Botran (2014). *A review of auto-scaling techniques for elastic applications in cloud environments.* Journal of Grid Computing.

Nations, D. (2018). What exactly is a web application? <https://www.lifewire.com/what-is-a-web-application-3486637>. Online, acessado em: 15/10/2018.

NIST (2011). *The nist definition of cloud computing.* <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>.



Figure 4. Utilização das instâncias no início dos testes



Figure 5. Utilização das instâncias no final dos testes