

HIPERFCLOUD: UM PROJETO DE ALTO DESEMPENHO EM NUVEM

RESUMO

Computação em nuvem é uma necessidade real para os ambientes de pesquisas e empresas. Embora bastante usada e estudada nos últimos anos, a computação em nuvem ainda traz diversos desafios. Um deles é a obtenção de alto desempenho, sendo o principal foco do projeto HiPerfCloud. Esta é uma tarefa complexa, pois é preciso lidar com diversos níveis de abstração, combinar tecnologias, avaliar modelos de implantação e usar soluções adequadas. Este artigo irá apresentar o projeto de pesquisa, seus principais objetivos e os principais resultados alcançados até o momento. Além disso, demonstrar as perspectivas da pesquisa no projeto.

Palavras Chaves: Computação em Nuvem. Alto Desempenho. Projeto de Pesquisa.

1 INTRODUÇÃO

A computação em nuvem é composta por modelos estruturados em uma pilha de três camadas. Na base, o modelo infraestrutura como serviço (IaaS) é responsável pelo gerenciamento da virtualização e do hardware, bem como oferecer a infraestrutura para as outras camadas (ex. instância e rede virtual). Logo acima, a plataforma como serviço (PaaS) oferece um ambiente implantado para atender as necessidades no desenvolvimento de uma aplicação específica (ex. banco de dados, linguagem de programação e sistema operacional). Já no topo da pilha, o software como serviço (SaaS) é o nível mais alto de abstração, oferecendo o produto final aos usuários (ex. editor de texto, multimídia e sistema de arquivos) [BUYYYA 2013].

A qualidade de serviço (QoS) é muito importante para a computação em nuvem, onde normalmente existe um acordo de nível de serviço (SLA) entre o provedor e o usuário. Um dos aspectos de QoS é o desempenho, sendo um dos principais desafios na nuvem devido as abstrações de software, combinações de tecnologias e modelos de implantação existentes. Esta foi a principal motivação para o início das pesquisas na camada de IaaS, a qual exerce grande influência no desempenho em todos os níveis da pilha [THOMÉ 2013] [MARON 2014] [VOGEL 2015].

O projeto HiPerfCloud¹ teve origem a partir de 2 Trabalhos de Conclusão do Curso (TCC) de redes de computadores da SETREM², realizados no Laboratório de Estudos Avançados para Computação em Nuvem (LARCC)³ e que estavam ligados à nossa motivação. O primeiro trabalho [THOMÉ 2013] mostrou que as ferramentas de administração em nuvem analisadas, possuem características específicas e necessitam de cenários de implantação exclusivos para atender os requisitos básicos da ferramenta. Assim, todos os resultados de [THOMÉ 2013] [MARON 2014a] serviram de inspiração para avaliar o desempenho dessas ferramentas e também para consolidar a escolha delas no segundo trabalho. Desta forma, o trabalho de [MARON 2014] implantou e avaliou o desempenho das ferramentas OpenStack e OpenNebula em um cenário de testes, considerando o desempenho de isolamento de recursos da infraestrutura (processador, memória, disco e rede) e de aplicações científicas (memória distribuída e compartilhada) [MARON 2014b] [MARON 2015].

Impulsionado pelos resultados dos trabalhos, o projeto HiPerfCloud tem como objetivo de pesquisa avaliar e analisar o desempenho de ferramentas, tecnologias e modelos de implantação envolvidos na camada de IaaS. O projeto também tem como objetivo trazer uma pesquisa de qualidade para a instituição, bem como levar conhecimento tecnológico para empresas e a comunidade. O grande ganho está no âmbito

¹ <http://hiperfcloud.setrem.com.br>

² <http://setrem.com.br>

³ <http://larcc.setrem.com.br>

pedagógico, pois aproxima alunos e professores com temas relacionados a computação em nuvem, motivando o estudo e a criação de novas tecnologias.

2 METODOLOGIA

Para atingir os objetivos de pesquisa, estudos exploratórios são realizados, fazendo um levantamento do estado da arte e encontrando pesquisas relacionadas. Este estudo é conduzido por alunos de iniciação científica e de conclusão de curso (dos cursos de graduação em redes e sistemas de informação), orientados por pesquisadores do LARCC. Depois, existe uma discussão em grupo, onde são delimitados os temas de pesquisa dentro do escopo do projeto. Embora cada aluno foque em um tema específico, todos os envolvidos colaboram e participam das discussões para resolver problemas e descobrir novas soluções. Na fase de desenvolvimento, são aplicadas técnicas de experimentos para coletar dados e interpretar os resultados, os quais serão disponibilizados para a comunidade através da publicação de artigos em conferências e revistas. Nos objetivos pedagógicos, o projeto busca o apoio de empresas para poder oferecer bolsas de iniciação científica. Com isso, os estudantes podem se dedicar exclusivamente as atividades de pesquisa, melhorando o seu conhecimento e também colaborando com a comunidade. Por outro lado, as empresas da região se beneficiam da localização do projeto (SETREM) para obter profissionais mais qualificados, e ainda aproveitam do conhecimento produzido na pesquisa para implantar soluções tecnológicas eficientes.

3 RESULTADO E DISCUSSÕES

Atualmente, o HiPerfCloud conta com uma equipe de oito pessoas (entre elas estão pesquisadores, colaboradores e alunos de iniciação científica). No contexto do projeto, foram produzidos 3 TCCs e publicados 6 artigos completos em conferências/revistas. Em parceria com a empresa Abase Sistemas⁴, o projeto oferece bolsa de iniciação científica desde 2014/2 e vem adquirindo hardware computacional para a realização das avaliações de desempenho (dentro do LARCC). Com os resultados da pesquisa, outros projetos em andamento na instituição (ex. Bioagropec e PowerDry) começaram a usar a infraestrutura adquirida através do HiPerfCloud.

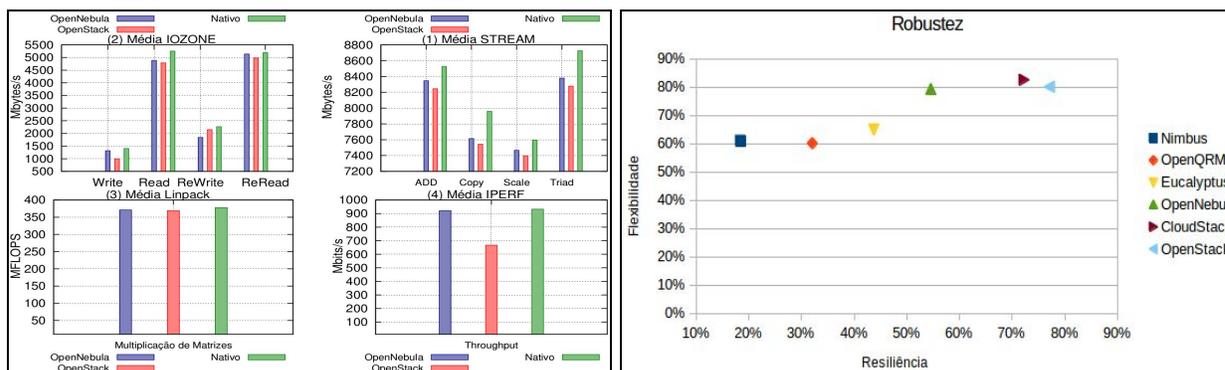
Os primeiros resultados de desempenho foram obtidos com o trabalho [MARON 2014]. A Figura 1(a) demonstra alguns destes obtidos em testes realizados na infraestrutura do projeto HiPerfCloud, que se encontram nas dependências do LARCC. Os gráficos representam os resultados de desempenho de memória RAM (1), armazenamento (2), processamento (3) e rede (4). A implantação foi com as ferramentas OpenStack e OpenNebula em *clusters* homogêneos. Os resultados com o ambiente nativo servem como referência para os obtidos em nuvem, na qual nota-se que existe uma diferença significativa. Acredita-se que o principal motivo para a diferença entre as ferramentas estão nos componentes existentes em sua estrutura. Por exemplo, o OpenStack contém um número maior de componentes internos para o gerenciamento dos recursos. Assim como pode ser visualizado no teste de rede, há uma perda significativa decorrente do método de implantação das redes e interfaces virtuais.

Os resultados obtidos na Figura 1(a), serviram como base para o entendimento individual de cada aspecto analisado (memória, rede, processador, armazenamento). Com isso, em uma outra etapa foram executados testes com aplicações científicas utilizando benchmarks sintéticos para avaliação de alto desempenho, que exploravam cargas de trabalho paralelas distribuídas através da suite MPI (*Message Passing Interface*) e local com a suite OpenMP (*Open Multi-processing*) [MARON 2015]. Os resultados em nuvem se mostram otimistas para execuções deste tipo de aplicações. Assim como nos resultados de isolamento de recursos da infraestrutura, OpenNebula se mostrou mais eficiente.

Apesar dos estudos anteriores de [THOMÉ 2013] [THOMÉ 2013a], não foram pesquisados aspectos relacionados a flexibilidade, resiliência e robustez. Diante disso, no trabalho de [VOGEL 2015] foram estudadas novas referências para conseguir mensurar de forma quantitativa estas variáveis. A Figura 1(b) apresenta de forma gráfica e escalar as ferramentas de código aberto para implantação de nuvens

⁴ <http://www.abase.com.br>

IaaS. Assim, é possível visualizar que existem contrastes relacionados com a robustez (flexibilidade e resiliência). Isso ocorre por que as ferramentas são desenvolvidas e implantadas com arquitetura, linguagens e propostas diferentes. A robustez impacta diretamente na capacidade de uma nuvem atender as demandas de clientes e aplicações (QoS). Os resultados das avaliações foram organizados com o valor percentual de 0 - 100%. Os resultados do *survey* mostraram que CloudStack e OpenStack são as mais robustas. CloudStack é mais flexível e OpenStack é mais resiliente. OpenNebula ficou na terceira posição em ambos os cenários. Por outro lado, ferramentas como Eucalyptus, OpenQRM e Nimbus ficaram com níveis baixos de robustez em função de aspectos particulares⁵.



a. Avaliação do Desempenho. b. Avaliação da Robustez.
 Figura 1: Principais descobertas nas pesquisas realizadas dentro do projeto HiPerfCloud.

4 CONCLUSÕES

Este artigo apresentou o projeto HiPerfCloud, contextualizando o cenário de pesquisa e discutindo alguns resultados obtidos através dele. Com os resultados de [MARON 2015] [MARON 2014] [MARON 2014a] [MARON 2014b], fica evidente que a nuvem é uma alternativa para a execução de aplicações com um desempenho comparável ao ambiente nativo. A principal descoberta para a comunidade é que as ferramentas podem de fato impactar no desempenho, umas menos e outras mais. Independente se as cargas são destinadas para o estresse de recursos específicos (isolamento de recursos da infraestrutura) ou cargas paralelas (local ou distribuída). Adicionalmente, traçou-se uma metodologia de experimentos para comparar estatisticamente as diferenças de desempenho. Posteriormente, a metodologia foi evoluída através da pesquisa de [VOGEL 2015], onde foi descoberta a forma mais justa de comparar se existem diferenças significativas de desempenho entre as ferramentas. Isso foi possível através das descobertas resultantes de um *survey* do estado da arte sobre o desempenho em ambiente virtuais.

Tendo em vista que cada virtualizador atende QoS em cargas específicas [BUYA 2013], a robustez encontrada também tem uma relação forte com o desempenho. Esta descoberta permite agora que os usuários tenham uma visão geral das ferramentas, podendo optar por aquela que se mostra mais apropriada dentro das necessidades do projeto de nuvem [VOGEL 2015]. Além disso, foi possível demarcar algumas carências a serem trabalhadas no futuro para que as ferramentas possam melhorar. Isso também motivou a pesquisa secundária de [ROVEDA 2015] [ROVEDA 2015a] para estudar mais de perto a interface de gerenciamento, revelando que OpenStack é atualmente mais completa neste aspecto quando comparado com OpenNebula e CloudStack.

As perspectivas de pesquisa para o HiPerfCloud são de: (I) comparar e avaliar o desempenho de todas as ferramentas open source de IaaS; (II) explorar novos tipos de cargas de trabalho através de outros *benchmarks* (ex. aplicações corporativas e de *Big Data*); (III) avaliar o impacto dos diferentes modelos de

⁵ Os aspectos particulares de cada uma delas estão detalhados em [VOGEL 2015], páginas 67 - 107.

implantação suportados pelas ferramentas; (IV) analisar o quanto os recursos de segurança podem impactar nelas; (V) avaliar e comparar o provisionamento e orquestração de instâncias. Paralelamente, o projeto pretende adquirir novos parceiros para ampliar o número de bolsas e recursos de estudo e pesquisa.

REFERÊNCIAS

[BUYA 2013] BUYA, Rajkumar; VECCHIOLA, Christian; SELVI, Thamarai. 2013. **Mastering Cloud Computing: Foundations and Applications Programming**. Elsevier Science. ISBN 9780124095397

[MARON 2014] MARON, Carlos.; GRIEBLER, Dalvan. 2014. **Avaliação e Comparação da Computação de Alto Desempenho em Ferramentas OpenSource de Administração de Nuvem Usando Estações De Trabalho**. - Trabalho de Conclusão do Curso Superior de Tecnologias em Redes de Computadores. Tecnologia em Redes de Computadores. Três de Maio: SETREM.

[MARON 2014a] MARON, Carlos; GRIEBLER, Dalvan; VOGEL, Adriano; SCHEPKE, Claudio. **Avaliação e Comparação do Desempenho das Ferramentas OpenStack e OpenNebula**. In: ESCOLA REGIONAL DE REDES DE COMPUTADORES (ERRC), 12., Canoas. Anais. Sociedade Brasileira de Computação, 2014. p.1–5.147

[MARON 2014b] MARON, Carlos; GRIEBLER, Dalvan; SCHEPKE, Claudio. **Comparação das Ferramentas OpenNebula e OpenStack em Nuvem Composta de Estações de Trabalho**. In: ESCOLA REGIONAL DE ALTO DESEMPENHO DO ESTADO DO RIO GRANDE DO SUL (ERAD/RS), 14., Alegrete, RS, Brazil. Anais. . . Sociedade Brasileira de Computação, 2014. p.173–176.

[MARON 2015] MARON, Carlos; GRIEBLER, Dalvan; SCHEPKE, Claudio. **Em Direção à Comparação do Desempenho das Aplicações Paralelas nas Ferramentas OpenStack e OpenNebula**. In: ESCOLA REGIONAL DE ALTO DESEMPENHO DO ESTADO DO RIO GRANDE DO SUL (ERAD/RS), 15. Gramado, RS, Brazil. Anais. Sociedade Brasileira de Computação, 2015.

[ROVEDA 2015] ROVEDA, Demetrius; VOGEL, Adriano; GRIEBLER, Dalvan. **Understanding, Discussing and Analyzing the OpenNebula and OpenStack's IaaS Management Layers**. Revista Eletrônica Argentina-Brasil de Tecnologias da Informação e da Comunicação, Três de Maio, Brazil, v.3, n.1, p.15, August 2015.

[ROVEDA 2015a] ROVEDA, Demetrius; VOGEL, Adriano; MARON, Carlos; GRIEBLER, Dalvan; SCHEPKE, Claudio. **Analisando a Camada de Gerenciamento das Ferramentas CloudStack e OpenStack para Nuvens Privadas**. In: ESCOLA REGIONAL DE REDES DE COMPUTADORES (ERRC), 12., Passo Fundo. Anais. Sociedade Brasileira de Computação, 2015.

[THOMÉ 2013] THOMÉ, Bruna; HENTGES, Eduardo; GRIEBLER, Dalvan. 2013. **Análise e Comparação de Ferramentas Open Source de Computação em Nuvem para o Modelo de Serviço IaaS** - Trabalho de Conclusão do Curso Superior de Tecnologias em Redes de Computadores. Tecnologia em Redes de Computadores. Três de Maio: SETREM.

[THOMÉ 2013a] THOMÉ, Bruna; HENTGES, Eduardo; GRIEBLER, Dalvan. **Computação em Nuvem: Análise Comparativa de Ferramentas Open Source para IaaS**. In: ESCOLA REGIONAL DE REDES DE COMPUTADORES (ERRC), 11., Porto Alegre, RS, Brazil. Anais. Sociedade Brasileira de Computação, 2013. p.4.

[VOGEL 2015] VOGEL, Adriano; GRIEBLER, Dalvan. 2015. **Surveying the Robustness and Analyzing the Performance Impact of Open Source Infrastructure as a Service Management Tools** - Trabalho de Conclusão do Curso Superior de Tecnologias em Redes de Computadores. Tecnologia em Redes de Computadores. Três de Maio: SETREM.