

Avaliação e Comparação do Desempenho das Ferramentas OpenStack e OpenNebula

Carlos A. F. Maron¹, Dalvan Griebler², Adriano Vogel¹, Claudio Schepke³

¹Curso Superior de Tecnologia em Redes de Computadores – Faculdade
Três de Maio (SETREM) – Três de Maio – RS – Brasil

²Programa de Pós-Graduação em Ciência da Computação – Pontifícia
Universidade Católica do Rio Grande do Sul (PUCRS) – Porto Alegre – RS – Brasil

³Laboratório de Estudos Avançados em Computação (LEA) - Campus Alegrete
Universidade Federal do Pampa (UNIPAMPA) – Alegrete – RS – Brasil

{francocaam, adrianovogel03}@gmail.com,
dalvan.griebler@acad.pucrs.br, claudioschepke@unipampa.edu.br

Resumo. *A computação em nuvem está cada vez mais presente nas infraestruturas corporativas. Por causa disso, diversas ferramentas estão sendo criadas para auxiliar na administração dos recursos na nuvem. O objetivo deste trabalho é avaliar o impacto que as ferramentas OpenStack e OpenNebula (implantadas em um ambiente de nuvem privado) podem causar no desempenho dos sistemas de memória, armazenamento, rede, e processador. Os resultados obtidos mostram que o desempenho no OpenStack é significativamente melhor nos testes do sistema de armazenamento, enquanto que no OpenNebula o restante dos testes foram melhores.*

Palavras-Chave: *Computação em Nuvem, Avaliação de Desempenho, OpenStack, OpenNebula.*

1. Introdução

A computação em nuvem é vista positivamente pelas empresas e instituições por oferecer recursos em diferentes níveis de serviço, como por exemplo, o modelo de serviço IaaS (*Infrastructure as a Service*), que permite a criação de máquinas virtuais na nuvem; SaaS (*Software as a Service*) que agrega *software* na nuvem; e PaaS (*Platform as a Service*) que fornece ambiente para desenvolvimento de *software* na nuvem. Destes 3 modelos básicos de serviços, a computação em nuvem pode ser implantada de modo privado, público, híbrido e comunitário [Marinescu 2013].

Um dos fatores questionáveis em ambientes de computação em nuvem é o desempenho, pois existe uma camada de virtualização operando abaixo dos serviços que são oferecidos aos usuários. Atualmente, existem diversas ferramentas que nos permitem implantar uma infraestrutura como serviço, e informações sobre o desempenho são de extrema importância para realizar futuras melhorias e escolher a ferramenta mais adequada para implantações. O objetivo deste trabalho é a avaliação de desempenho das ferramentas do modelo de serviço IaaS (OpenStack e OpenNebula) e a comparação entre elas. Embora a literatura nos contemple com trabalhos que avaliam o desempenho dos principais virtualizadores [Hwang et al. 2013, Xavier et al. 2013], é preciso verificar se as ferramentas de administração de nuvem influenciam no desempenho, pois elas são uma camada de abstração para lidar com recursos de rede e periféricos virtualizados.

A necessidade de avaliação do desempenho vem sendo notada em trabalhos realizados anteriormente [Thome et al. 2013, Maron et al. 2014]. Nestes estudos, identificou-se que cada ferramenta é implementada com tecnologias e métodos particulares, além de oferecer diferentes funcionalidades para lidar com a camada de virtualização. Essa diversidade pode afetar positivamente ou negativamente o desempenho da ferramenta. Isso nos motivou a avaliar se de fato existem diferenças entre as ferramentas, e entre o ambiente nativo. As principais contribuições deste trabalho são elencadas a seguir:

- Avaliação do desempenho das ferramentas de administração de nuvem OpenStack e OpenNebula.
- Análise comparativa das diferenças de desempenho entre as ferramentas e o ambiente nativo.

O artigo está organizado em 5 seções. Na Seção 2 são descritos os principais conceitos de computação em nuvem e aspectos sobre ferramentas de administração. A Seção 3 discute os estudos relacionados com a presente pesquisa. Na Seção 4, é descrita a metodologia que envolve os processos de instalação e testes realizados. Ao final, na Seção 5 são apresentados os resultados obtidos com as ferramentas OpenStack e OpenNebula.

2. Computação em Nuvem

O modelo da computação em nuvem baseia-se na entrega de serviços como infraestrutura, ou softwares para usuários que não possuem esses recursos em suas dependências, ou necessitam temporariamente. O baixo investimento financeiro é um dos principais motivos para o uso destes serviços, pois as modalidades de cobrança de provedores, baseiam-se no *pay-per-use* (pague pelo uso) [Marinescu 2013].

Na computação em nuvem privada, a implantação é feita no ambiente interno da empresa, evitando assim, a dependência de provedores externos (Nuvem Pública) e de eventuais atrasos na rede fora do domínio da empresa (Internet). A nuvem privada oferece um aumento nos níveis de segurança e privacidade, e acaba sendo atrativa para as empresas, pois oferece também elasticidade, flexibilidade, economia de energia, capacidade de customização e agilidade nos negócios. Além disso, oferece benefícios para ambientes científicos. No entanto, causa confusão quando comparada com ferramentas de virtualização, pois alguns *data centers* as utilizam antes mesmo do conceito ter surgido.

De modo diferente, a computação em nuvem é baseada na oferta de recursos e serviços (IaaS, SaaS, PaaS), o que faz necessário uma camada de abstração (ferramentas de administração) sobre as tecnologias de virtualização para facilitar o gerenciamento destes serviços [Marinescu 2013]. Neste trabalho, optou-se pelo virtualizador KVM (*Kernel-based Virtual Machine*), que se mostrou mais compatível entre as ferramentas. Além disso, a forma com que KVM está inserido no kernel do Linux possibilita um ganho de desempenho em relação à outros virtualizadores *open source*.

2.1. Ferramentas de Administração

As ferramentas de administração de nuvens possuem diferentes características de implementação e buscam atender necessidades específicas aos usuários. Isso se tornou evidente no trabalho de [Thome et al. 2013], onde foram avaliadas e comparadas as características operacionais de várias delas. A seguir, é realizada uma breve descrição sobre as ferramentas utilizadas nesta pesquisa.

O OpenStack nasceu de um projeto da Rackspace e da NASA. Atualmente é usado para construção de nuvens IaaS públicas e privadas. A ferramenta trabalha com um conjunto de componentes, onde é possível implementar redes virtuais (Neutron), discos virtuais no formato LVM (*Cinder*), controlar a camada de virtualização em conjunto com os virtualizadores (Nova), e oferecer serviços para bancos de dados (Trove) [OpenStack 2014b]. Ainda são necessários componentes e serviços de terceiros para contribuir nas atividades da ferramenta, como o OpenVSwitch para implementar interfaces virtuais e, o *middleware* RabbitMQ ou ZeroMQ para troca de mensagens entre serviços.

OpenNebula surgiu em 2008, também como uma ferramenta (*opensource*) para ambientes de nuvem privada, pública e híbrida. É implantada baseando-se nos modelos clássicos de *cluster*. Portanto, existe um *node* da infraestrutura que é usado como *master/front-end*, o qual executa os principais serviços da ferramenta e gerencia o restante dos *nodes* que são os escravos. Além do *core* principal da ferramenta OpenNebula (ONED) que controla grande parte dos componentes, existe o OneGate e OneFlow [OpenNebula 2014]. A comunicação entre os *nodes* é feita através do protocolo SSH e, através do recurso NFS (*Network File System*), a ferramenta permite o provisionamento das instâncias (Máquinas Virtuais) na infraestrutura [OpenNebula 2014].

3. Trabalhos Relacionados

Olhando para a literatura, nota-se que a comunidade tem-se preocupado com diversas questões relacionadas ao desempenho de computação em nuvem. Como por exemplo, a pesquisa de [Corradi et al. 2014] destaca a questão da degradação do desempenho quando um grande número de máquinas virtuais está executando em um *host*. O resultado conclusivo da pesquisa mostrou que quando o processamento das máquinas virtuais alcançam níveis próximos a 100%, a degradação real do desempenho é notável e mesmo assim o consumo de energia é eficiente. Estes resultados foram obtidos em um ambiente usando: OpenStack (versão Diablo) como ferramentas de administração da nuvem, KVM como virtualizador base, e *benchmarks* MapReduce e IPerf para a análise da rede.

Por outro lado, a literatura também nos contempla com trabalhos como o de [Xavier et al. 2013] e [Hwang et al. 2013], que se atentam para a avaliação do desempenho dos principais virtualizadores para computação em nuvem. O foco do trabalho de [Xavier et al. 2013] foi a virtualização em *container* (virtualização a nível de sistema operacional) com os seguintes virtualizadores: Linux Vserver, OpenVZ e LXC. Os testes foram realizados com os *benchmarks*: STREAM (teste de memória), IOzone (teste de disco), NetPipe (avaliação de rede). Complementando, na pesquisa de [Hwang et al. 2013] foram avaliados os virtualizadores Hyper-V, VMware, KVM e XEN. Os testes foram efetuados com os seguintes *benchmarks*: ByteMark (teste de Processador), RAMSpeed (teste de memória), Bonnie++ e Filebench (teste de Disco).

Os resultados de [Xavier et al. 2013] mostram proximidade no desempenho entre virtualizadores e ambiente nativo, afirmando que é necessário apresentar melhores resultados para execução de aplicações de alto desempenho. Já [Hwang et al. 2013], destaca que os resultados não foram satisfatórios com os virtualizadores baseados em *hypervisor*. Além disso, este relata haver dificuldades para a definição destas tecnologias de virtualização e aconselha uma análise específica de todo o cenário para auxiliar na escolha, pois o desempenho entre elas foi diferente.

Assim como os trabalhos relacionados, esta pesquisa contribui para avaliar o desempenho da computação em nuvem. Diferente do que é visto nos trabalhos elencados, o foco está na avaliação do desempenho das ferramentas de administração. Isso porque, acredita-se que devido a diversidade de implementação entre as ferramentas evidenciado em [Thome et al. 2013, Maron et al. 2014], o desempenho da nuvem pode ser afetado de forma positiva ou negativa.

4. Metodologia

Para avaliar o desempenho das ferramentas, foram usados *benchmarks* para analisar os componentes da infraestrutura virtual (virtualizador KVM) e Nativa, e comparar as diferenças no desempenho. Sendo assim, os seguintes *benchmarks* foram executados com a configuração padrão: IPerf (avaliação da rede, coletando somente os dados de *throughput*); IOzone (teste das unidades de armazenamento, analisando *Read*, *Reread*, *Write*, *Rewrite* aplicado à um arquivo de 100 MB); STREAM (teste da memória, analisando as funções *ADD*, *SCALE*, *TRIAD*, *COPY*); LINPACK (teste do processador, usando matriz de 4.000x4.000 para executar cálculos de pontos flutuantes). Cada *benchmark* foi executado 40 vezes para o cálculo da média.

A infraestrutura utilizada é composta por 8 computadores idênticos com Intel Core i5 650 - 3.20 GHz, memória RAM de 4 GB DDR3 de 1333 MHz, disco de 500 GB operando em sata II, e todos os *nodes* operando em rede 10/100 Mbits. O sistema operacional usado em todas as instalações foi o Ubuntu Server 12.04. A versão de OpenStack foi a Havana e para OpenNebula foi a 4.7.80. Cada nuvem foi isoladamente implantada com 4 computadores, sendo 1 *front-end* e 3 *hosts*. No ambiente virtual, as instâncias foram configuradas de acordo com a capacidade total do *host* físico (referente a memória, disco e processadores).

5. Resultados

A Figura 1 apresenta os resultados obtidos de largura de banda de memória, rede, acesso a disco e performance de processamento. Os gráficos da Figura 1(a).1 demonstram o desempenho alcançado pelo *benchmark* STREAM em cada um dos ambientes. O ambiente OpenNebula apresentou melhores resultados que OpenStack em termos de largura de banda de memória. Os resultados dos testes utilizando o algoritmo Linpack são mostrados na Figura 1(a).3. Novamente, nota-se que o desempenho no OpenStack é inferior ao OpenNebula. Isso ocorre porque o componente *nova-compute-kvm* realiza requisições ao banco de dados em tempo de execução, a fim de armazenar informações sobre a máquina virtual. Esta possibilidade de perda é também reconhecida pelo OpenStack. Além disso, este componente pode oferecer riscos a segurança do ambiente [OpenStack 2014a].

Na Figura 1(a).2 são destacados os resultados obtidos nos testes das funções básicas de disco. O ambiente Nativo acompanhou em certo grau o desempenho nas funções de reescrita (*REWRITE*) e releitura (*REREAD*), em relação ao *WRITE* e *READ* simples, sendo um resultado previsível devido ao uso de *cache* nas operações de disco. A ferramenta OpenStack acompanhou esse ganho nessas funções. Embora ambas as ferramentas utilizem o mesmo formato de disco (LVM), OpenNebula não acompanhou o ganho nas funções *REWRITE* e *REREAD*. Isso acontece porque, nesta ferramenta, os discos se encontram em uma unidade compartilhada pela rede (usando NFS). Além disso, o desempenho é prejudicado pela baixa largura de banda (rede *Megabit*).

Conforme Figura 1(a).4, o bom desempenho do *benchmark* de rede na ferramenta OpenNebula pode ser justificado pelo *driver* básico da ferramenta (Dummy) que não impõe políticas de tráfego na rede virtual. Em OpenStack é utilizado Neutron e OpenvSwitch, os quais provocam dois saltos a mais na rede.

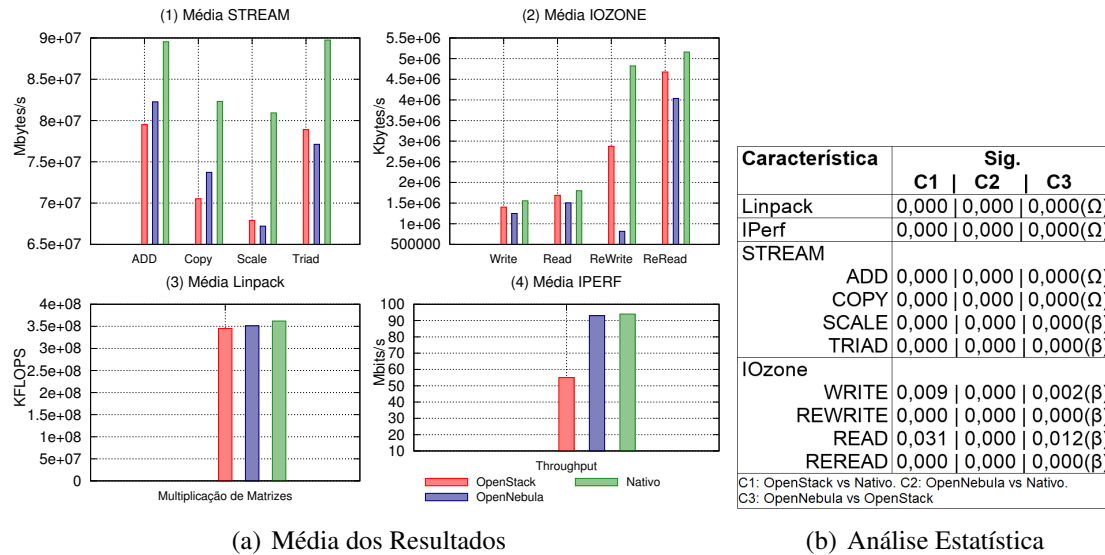


Figura 1. Análise dos resultados

Com o auxílio do software SPSS [Field 2009], foram realizados os testes estatísticos para identificar se as diferenças de desempenho foram significativas com uma confiança de 95%. Portanto, para considerar que as médias aritméticas sejam significativamente diferentes, a variável Sig. deve ser menor que 0,05. Os resultados estão ilustrados na Figura 1(b), onde o símbolo Ω indica a superioridade do OpenNebula e o β do OpenStack. Assim, é notório que o desempenho nos ambientes são significativamente diferentes e o OpenNebula possui vantagem na comparação entre as ferramentas.

6. Conclusão

Este trabalho teve como objetivo avaliar o desempenho das ferramentas de administração de nuvens (OpenStack e OpenNebula) e comparar as diferenças para identificar o impacto que elas podem causar. Sucintamente, os resultados mostraram que as diferenças de desempenho são significativas e que cada ferramenta é melhor em determinados aspectos. Para um ambiente controlado, concluímos que OpenStack é recomendado somente para aplicações que utilizam exaustivamente disco enquanto OpenNebula é melhor em aplicações que usam rede, processador e memória. No entanto, é importante destacar que o desempenho em relação ao ambiente nativo ainda está muito distante, algo que [Zia and Khan 2012] já havia relatado. Portanto, mesmo que OpenNebula consiga administrar melhor a maior parte dos recursos de uma nuvem, é preciso que os virtualizadores e as ferramentas progridam em conjunto, para atingirem desempenhos mais próximos do ambiente nativo.

Embora o foco deste artigo são os resultados de desempenho, antes de executar os experimentos foi necessária a implantação do ambiente. Vale relatar esta experiência nesta etapa da pesquisa, pois a instalação e configuração exigiram um esforço extra. Percebe-se

uma dificuldade maior em instalar OpenStack (versão Folsom) em relação a OpenNebula (versão 4.2). Em [Maron et al. 2014], as mesmas dificuldades foram encontradas (para as versões OpenStack Havana e OpenNebula 4.7.80), devido a quantidade de componentes e arquivos que necessitam ser configurados em OpenStack. Além disso, notou-se a instabilidade de alguns dos serviços (Ex. Neutron e Nova-Compute) providos por OpenStack. Ambos os aspectos podem influenciar na escolha da implantação de uma nuvem.

Os próximos passos da pesquisa são: (I) avaliar o desempenho de aplicações científicas e corporativas nos ambientes desta pesquisa, (II) avaliar o desempenho das ferramentas utilizando outros virtualizadores, (III) implantar e testar outras ferramentas de administração de nuvem para comparar com as aqui avaliadas.

Referências

- Corradi, A., Fanelli, M., and Foschini, L. (2014). VM consolidation: A Real Case Based on OpenStack Cloud. *Future Generation Computer Systems*, 32(0):118 – 127.
- Field, A. (2009). *Discovering Statistics Using SPSS*. SAGE, Dubai, EAU.
- Hwang, J., Zeng, S., Wood, T., et al. (2013). A Component-Based Performance Comparison of Four Hypervisors. In *Integrated Network Management (IM 2013), 2013 IFIP/IEEE International Symposium on*, pages 269–276. IEEE.
- Marinescu, D. (2013). *Cloud Computing: Theory and Practice*. Elsevier, 225 Wyman Street, Waltham, 02451, USA.
- Maron, C. A. F., Griebler, D., and Schepke, C. (2014). Comparação das Ferramentas OpenNebula e OpenStack em Nuvem Composta de Estações de Trabalho. In *14ª Escola Regional de Alto Desempenho do Estado do Rio Grande do Sul - ERAD/RS*, pages 173–176, Alegrete, RS, Brazil. Sociedade Brasileira de Computação - SBC.
- OpenNebula (2014). OpenNebula (Official Page) <<http://opennebula.org/>>. Last access in October, 2014.
- OpenStack (2014a). Designing for Cloud Controllers and Cloud Management <http://docs.openstack.org/openstack-ops/content/cloud_controller_design.html>. Last access in November, 2014.
- OpenStack (2014b). OpenStack roadmap <<http://openstack.org/software/roadmap/>>. Last access May, 2014.
- Thome, B., Hentges, E., and Griebler, D. (2013). Análise e Comparação de Ferramentas Open Source de Computação em Nuvem para o Modelo de Serviço IaaS. In *Escola Regional de Redes de Computadores (ERRC)*, pages 1–4, Porto Alegre, RS, Brazil. Sociedade Brasileira de Computação - SBC.
- Xavier, M., Neves, M., Rossi, F., Ferreto, T., Lange, T., and Rose, C. D. (2013). Performance Evaluation of Container-based Virtualization for High Performance Computing Environments. In *21st Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, pages 233–240, Belfast, UK. IEEE.
- Zia, A. and Khan, M. (2012). Identifying Key Challenges in Performance Issues in Cloud Computing. In *Modern Education and Computer Science Press*, pages 59–68. Modern Education and Computer - I.J.