

Medindo o Desempenho de Implantações de OpenStack, CloudStack e OpenNebula em Aplicações Científicas

Adriano Vogel¹, Carlos A. F. Maron², Dalvan Griebler^{1,2} Claudio Schepke³

¹ Laboratório de Pesquisas Avançadas para Computação em Nuvem (LARCC)
Faculdade Três de Maio (SETREM) – Três de Maio – RS – Brasil

² Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
Porto Alegre – RS – Brasil

³ Universidade Federal do Pampa (UNIPAMPA), Laboratório de Estudos Avançados (LEA),
Alegrete – RS – Brasil

adrianovogel103@gmail.com, carlos.maron@acad.pucrs.br
dalvan.griebler@acad.pucrs.br, claudioschepke@unipampa.edu.br

Resumo. *Ambientes de nuvem possibilitam a execução de aplicações sob demanda e são uma alternativa para aplicações científicas. O desempenho é um dos principais desafios, devido ao uso da virtualização que induz perdas e variações. O objetivo do trabalho foi implantar ambientes de nuvem privada com diferentes ferramentas de IaaS, medindo o desempenho de aplicações paralelas. Consequentemente, os resultados apresentaram poucos contrastes.*

1. Introdução

Computação em nuvem é um paradigma emergente que oferece recursos computacionais como serviços. O serviço mais popular e importante é o de IaaS (Infraestrutura como Serviço) tanto em provedores de nuvem pública quanto em nuvens privadas (para o domínio exclusivo de uma instituição). A relevância do serviço de IaaS é resultado dos demais serviços de nuvem (PaaS, SaaS, entre outros) serem camadas de *software* executadas em ambientes IaaS, o que causa uma dependência por poder computacional, robustez e eficiência [Iosup et al. 2011, Buyya et al. 2013, Vogel et al. 2016].

Para garantir QoS (*Quality of Service*) em ambientes de nuvem, o desempenho e eficiência na utilização dos recursos são vitais. Porém, a utilização da virtualização nestes ambientes causa alguns tipos de perdas no desempenho das aplicações. Além disso, para análise de QoS, as medições de desempenho das execuções são complexas e exaustivas. Dessa forma, esse são fatores que motivam pesquisas com o desempenho de aplicações.

Nesse estudo, cargas de trabalho renomadas são executadas em ambientes de nuvem privada. Essa avaliação continua estudos anteriores que introduziram medições nas ferramentas OpenStack e OpenNebula em outros ambientes computacionais. Por exemplo, o trabalho [Maron et al. 2014] fez uma introdução ao desempenho da infraestrutura e o [Maron et al. 2015] avaliou a eficiência e *speedup* com alguns *kernels* do NAS usando estações de trabalho. Por outro lado, esta abordagem apresenta uma nova implantação e avaliação de instâncias oferecidas por três implementações de nuvem. Além disso, um número maior de *kernels* foi executado com resultados das médias de tempo em um ambiente computacional mais robusto.

Esse trabalho é dividido em 4 seções. A Seção 2 apresenta os estudos relacionados. Na Seção 3 são mostrados os resultados dos experimentos nas implantações de nuvem. A última parte apresenta a conclusão e os trabalhos futuros desse estudo.

2. Trabalhos Relacionados

A opção de utilizar computação em nuvem para HPC vem recebendo uma crescente atenção da literatura. O estudo de [Akioka and Muraoka 2010] e [Jackson et al. 2010] analisam o desempenho de instâncias da EC2, usando o HPL (*High-Performance Linpack*) e experimentos do NAS. Os resultados de desempenho na nuvem comparados com máquinas físicas mostram perdas de desempenho, devido ao compartilhamento de recursos e pelo uso de virtualização, principalmente nas aplicações distribuídas que dependem do desempenho de rede. Por outro lado, o estudo de [Krishnan et al. 2014] analisou o desempenho de 3 arquiteturas para HPC usando uma aplicação de previsão do tempo. O primeiro ambiente foi baseado em um tradicional *cluster Beowulf*, o segundo ambiente foi de nuvem pública com instâncias oferecidas pela AWS e o terceiro ambiente foi um *cluster* virtual usando o KVM. O estudo conclui que ambientes de nuvem podem ser eficientes e ter ganhos de desempenho pela fácil escalabilidade.

Os trabalhos apresentam uma visão para desempenho em ambientes de nuvem e nativos. Nesse estudo o desempenho também é analisado usando aplicações consolidadas (NAS) utilizadas em alguns trabalhos relacionados. Por outro lado, enquanto a maioria dos trabalhos analisam ambientes de nuvens públicas, supercomputadores ou *clusters* nativos, em nosso estudo as aplicações são executadas em nuvem privada. Ainda, o tempo de execução é considerado para a análise dos resultados, algo não analisado nos trabalhos relacionados. Esse estudo também contrasta com os relacionados pois as execuções são realizadas em nuvens privadas, possibilitando a utilização de todo o poder computacional das máquinas físicas, algo impossível em instâncias de nuvem pública.

3. Resultados

Para a execução de experimentos, três *clusters* isolados foram criados usando 4 máquinas idênticas, cada uma com 24 GB de RAM (1333 MHz), processador Intel Xeon X5560 (quad-core 2.80GHz), discos SATA II (7200 RPM) e conectados em uma rede Giga-bit (10/1000). No ambiente nativo e nas instâncias foi utilizado o sistema operacional Ubuntu *Server* 14.04 (kernel 3.19.0). Optou-se por implementar os ambientes usando virtualização completa através do KVM (versão 2.0.0) e as ferramentas de nuvem OpenStack versão Kilo, OpenNebula 4.12 e CloudStack 4.5.2.

A arquitetura das implantações foi similar. Usando as diferentes ferramentas de IaaS, cada ambiente teve um servidor agindo como gerente de nuvem, o qual controla a infraestrutura (recursos e usuários) e aloca as instâncias nos nodos. Controlados pelo gerente de nuvem, o virtualizador e agente de nuvem foram instalados em cada servidor (nodo) que ofereciam os recursos virtuais para as instâncias de nuvem. Ambas as implantações utilizaram os volumes distribuídos pela rede, OpenNebula e CloudStack usando o protocolo NFS (Network File System) e imagens QCOW (*QEMU Copy On Write*) e OpenStack usou o protocolo iSCSI (Internet Small Computer System Interface) e imagens LVM (*Logical Volume Management*). Em ambas as implantações a conectividade entre as instâncias foi usando as *Bridges Linux* e o *driver VirtIO*. Para aumentar

o poder de processamento, em cada servidor foi alocada apenas uma instância que tinha todos os recursos disponíveis.

Para os testes de desempenho, o NPB-3.3 foi executado em MPI (*Message-Passing Interface*) com até 16 processos (usando 2 máquinas) e OMP (*Open Multi-Processing*) até 8 threads. Esse conjunto de experimentos é consolidado para testes de poder computacional. Foi utilizada a classe B com os kernels BT, FT, IS, MG, CG, EP, LU e SP os quais foram repetidos 40 vezes em cada ambiente. Na Figura 1 são mostrados os resultados dos kernels OMP. A Figura 2 apresenta os resultados do MPI.

De modo geral, os resultados dos kernels NPB-OMP apresentam tempos de execução semelhantes. Percebe-se alguns contrastes nas médias, principalmente, quando o número de processos é maior, o que também aumenta a variação devido aos múltiplos processos criados nos servidores. Nas execuções com NPB-OMP CG e IS fica mais evidente a concorrência ao canal de memória devido as características da aplicação.

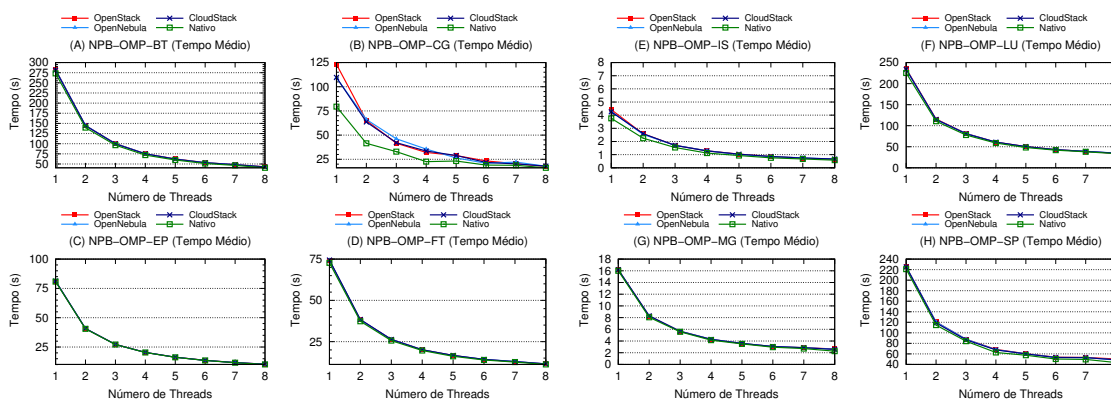


Figura 1. NAS OMP tempo de execução.

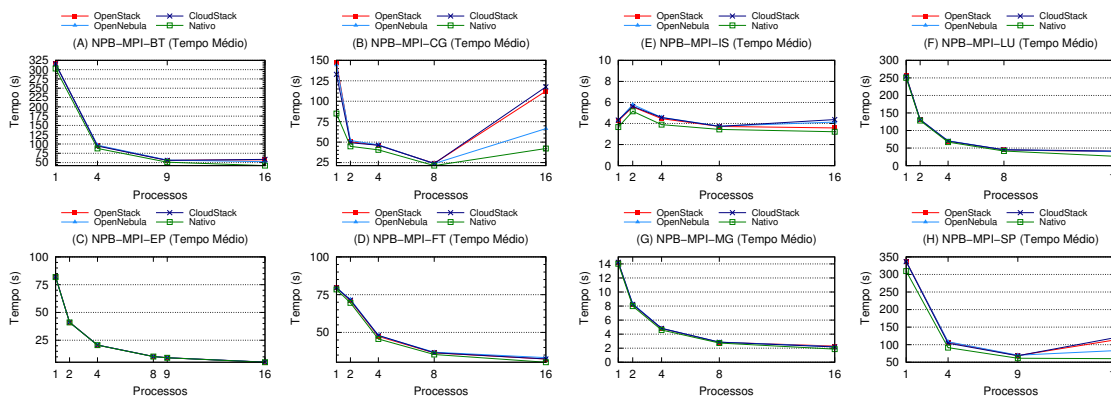


Figura 2. NAS MPI tempo de execução.

As médias dos tempos de execução do MPI nas instâncias em ambiente distintos foram próximas, o ambiente nativo teve ganhos com um número maior de processos (até 16), porém o *overhead* nos ambientes virtualizados foi pequeno. Entre as ferramentas, as médias de MPI-CG com 16 processos foram melhores nas instâncias da ferramenta OpenNebula, o mesmo se repete com o kernel SP. De modo geral, alguns tempos de execução diferentes sugerem um impacto causado pelo experimento ao invés dos ambientes. Outro

aspecto relevante observado é que em experimentos que usam intensivamente a rede tiveram algum tipo de perda nos ambientes virtualizados, o que é relacionado com os desafios da comunicação inter-processos, que depende principalmente de baixas latências de rede.

4. Conclusões

Este estudo apresentou uma avaliação de instâncias de ambientes de nuvem privada através da execução de experimentos do NAS. As médias dos tempos de execução em diferentes implantações foram similares na maioria dos *kernels*, o que indica que ambas as ferramentas podem oferecer recursos computacionais de forma eficiente e competitiva para aplicações científicas. Além disso, os resultados atingidos eram esperados, pois tecnologias iguais foram propositalmente implantadas para verificar se haveriam discrepâncias. Assim, confirma-se que não existem influências no desempenho durante a execução de cargas de trabalho nas instâncias de uma nuvem privada. Os experimentos realizados nos trabalhos anteriores apontaram diferenças entre as ferramentas [Maron et al. 2014] [Maron et al. 2015], constatando assim que embora fora usado o mesmo virtualizador, tal resultado foi em função do uso de tecnologias diferentes na implantação.

Os trabalhos futuros serão a customização da infraestrutura (virtualizadores, tecnologias de armazenamento e rede) das implantações para testar e comparar o desempenho das funções de gerenciamento das ferramentas (i.e., orquestração, alta disponibilidade e escalonamento).

Referências

- [Akioka and Muraoka 2010] Akioka, S. and Muraoka, Y. (2010). HPC benchmarks on Amazon EC2. In *Advanced Information Networking and Applications Workshops (WAINA), 2010 IEEE 24th International Conference on*, pages 1029–1034.
- [Buyya et al. 2013] Buyya, R., Vecchiola, C., and Selvi, S. (2013). *Mastering Cloud Computing*. McGraw Hill.
- [Iosup et al. 2011] Iosup, A., Yigitbasi, N., and Epema, D. (2011). On the performance variability of production cloud services. In *Cluster, Cloud and Grid Computing (CCGrid), 2011 11th IEEE/ACM International Symposium on*, pages 104–113.
- [Jackson et al. 2010] Jackson, K. R., Ramakrishnan, L., Muriki, K., Canon, S., Cholia, S., Shalf, J., Wasserman, H. J., and Wright, N. J. (2010). Performance analysis of high performance computing applications on the amazon web services cloud. In *Cloud Computing Technology and Science (CloudCom), 2010 IEEE Second International Conference on*, pages 159–168.
- [Krishnan et al. 2014] Krishnan, S., Veeravalli, B., Krishna, V. H., and Sheng, W. C. (2014). Performance Characterisation and Evaluation of WRF Model on Cloud and HPC Architectures. In *High Performance Computing and Communications (HPCC, CSS, ICSS)*, pages 1280–1287.
- [Maron et al. 2014] Maron, C. A. F., Griebler, D., Vogel, A., and Schepke, C. (2014). Avaliação e Comparação do Desempenho das Ferramentas OpenStack e OpenNebula. In *12th Escola Regional de Redes de Computadores (ERRC)*, pages 1–5, Canoas. Sociedade Brasileira de Computação.
- [Maron et al. 2015] Maron, C. A. F., Griebler, D., Vogel, A., and Schepke, C. (2015). Em Direção à Comparação do Desempenho das Aplicações Paralelas nas Ferramentas OpenStack e OpenNebula. In *15th Escola Regional de Alto Desempenho do Estado do Rio Grande do Sul (ERAD/RS)*, Gramado, RS, Brazil. Sociedade Brasileira de Computação.
- [Vogel et al. 2016] Vogel, A., Griebler, D., Maron, C. A. F., Schepke, C., and Fernandes, L. G. L. (2016). Private IaaS Clouds: A Comparative Analysis of OpenNebula, CloudStack and OpenStack. In *24rd Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, Heraklion Crete, Greece. IEEE.