

Análise de Viabilidade do Perfilamento de Aplicações de HPC Baseada em Contadores de *Hardware* na AWS

Luiz Fernando Althoff¹, Vanderlei Munhoz¹, Márcio Castro¹

¹Laboratório de Pesquisa em Sistemas Distribuídos (LaPeSD)
Universidade Federal de Santa Catarina (UFSC) - Florianópolis/SC

luiz.fernando.althoff@grad.ufsc.br, vanderlei.munhoz@posgrad.ufsc.br,
marcio.castro@ufsc.br

Resumo. Este artigo apresenta um estudo preliminar sobre a viabilidade do uso da técnica de perfilamento de aplicações de HPC baseada em contadores de hardware na AWS. O objetivo é entender as limitações e os impactos dessa técnica quando aplicada em um ambiente virtualizado. Os resultados evidenciam algumas limitações da AWS e uma pequena variação nos resultados quando a virtualização é utilizada.

1. Introdução

A Computação em Nuvem pode ser definida como “a entrega sob demanda de recurso computacional com o mínimo de esforço ou interação” [Mell et al. 2011]. Esse modelo de negócio democratizou o acesso a grandes infraestruturas de TI para milhões de pessoas e empresas, permitindo com que uma grande quantidade de serviços e produtos possam ser disponibilizados ao público de forma rápida e com baixo custo [Varghese and Buyya 2018].

As nuvens computacionais podem ser uma opção atrativa também para pesquisadores da área de Computação de Alto Desempenho (*High Performance Computing* – HPC), pois permite que grupos de pesquisa menores, com acesso limitado a *clusters* de HPC, possam fazer experimentos em larga escala de forma ágil e com custo muito inferior ao necessário para compra, gestão e manutenção de infraestruturas físicas *on-premise* [Saini and Sainis 2021]. Com intuito de facilitar a construção e configuração de *clusters* de HPC em nuvens computacionais de forma agnóstica de provedor, foi desenvolvida a ferramenta *open-source* denominada HPC@Cloud [Munhoz and Castro 2022]. A ferramenta automatiza todo o processo de alocação e configuração da infraestrutura, permitindo que os pesquisadores possam executar suas aplicações de HPC na nuvem sem se preocuparem com questões técnicas que envolvem a criação do ambiente.

Atualmente, a ferramenta HPC@Cloud não possui nenhum módulo de perfilamento de aplicações. Este módulo seria de grande valia para os pesquisadores da área de HPC, pois forneceria informações importantes sobre a execução das aplicações na nuvem e permitiria que otimizações específicas para a nuvem pudessem ser implementadas. As informações coletadas poderiam, por exemplo, ser utilizadas para a escolha dos tipos de máquinas a serem alocadas na nuvem que tivessem um melhor custo-benefício em função das características da carga de trabalho a ser executada.

Neste artigo é apresentado um estudo preliminar sobre a viabilidade do uso da técnica de perfilamento de aplicações de HPC baseada em contadores de *hardware* na AWS. Algumas métricas de desempenho foram coletadas durante a execução de aplicações do *NAS Parallel Benchmarks* [Bailey et al. 1991] em máquinas padrão (sem virtualização) e virtualizadas na AWS. Os resultados mostram que a técnica de perfilamento baseada em contadores de *hardware* só pode ser utilizada em um determinado

subconjunto de máquinas da AWS e que existe um pequeno impacto da virtualização nos resultados coletados.

Este artigo está organizado da seguinte forma. Na Seção 2 é apresentada, de forma sucinta, a ferramenta HPC@Cloud e é descrito como foi adicionado o suporte ao perfilamento de aplicações. Na Seção 3 é apresentada uma análise comparativa dos resultados obtidos com as aplicações do NAS quando executadas em máquinas virtualizada e *bare-metal* na AWS. Por fim, na Seção 4 são apresentadas as conclusões e discutidas algumas possibilidades de trabalhos futuros.

2. Perfilamento na Ferramenta HPC@Cloud

A ferramenta HPC@Cloud [Munhoz and Castro 2022] é um *software* de código aberto¹ que está sendo desenvolvido pelo Laboratório de Pesquisa em Sistemas Distribuídos (LaPeSD) da Universidade Federal de Santa Catarina (UFSC) para facilitar a construção de *clusters* para HPC em nuvens computacionais públicas. Atualmente, a ferramenta suporta a execução de aplicações paralelas implementadas em OpenMP e MPI² nos provedores Amazon Web Services (AWS) e Vultr Cloud, incluindo a capacidade de construir *clusters* do tipo *spot* na AWS [Munhoz and Castro 2022]. Além da interface, a ferramenta também disponibiliza imagens e *snapshots* de máquinas virtuais previamente construídas para acelerar a construção da infraestrutura, reduzindo-se assim os custos. Por fim, a ferramenta também possui suporte a criação e gestão de instâncias do tipo *spot* da AWS.

As ferramentas de perfilamento que fazem uso de contadores de *hardware* extraem informações diretamente da Unidade de Monitoramento de Desempenho (*Performance Monitoring Unit* – PMU) disponível nos processadores atuais. Uma das ferramentas mais utilizadas para coletar essas informações é a Perf³. A ferramenta Perf permite a coleta de diversas métricas relacionadas às unidades lógicas e aritméticas do processador, à memória *cache*, ao barramento, à memória RAM e entre outras. Neste trabalho, os arquivos de configuração da ferramenta HPC@Cloud foram modificados para incluir a instalação dos pacotes necessários para o funcionamento do Perf. Desta forma, os *clusters* criados pela ferramenta HPC@Cloud já possuirão o ambiente pronto para capturar as métricas das aplicações.

Durante a implementação do suporte ao perfilamento na ferramenta HPC@Cloud foi detectada uma limitação na AWS quanto ao acesso ao PMU em instâncias virtualizadas. Com a virtualização, há um suporte existente no nível do *hypervisor* para monitorar as máquinas virtuais (vPMU). Todavia, esse suporte está habilitado somente em instâncias da AWS que utilizem processadores Intel, e que tenham quantidade de vCPUs igual ao número de núcleos físicos do processador utilizado na máquina virtual.

3. Resultados Experimentais

Nesta seção é apresentada uma avaliação preliminar das limitações e impactos do uso da técnica de perfilamento baseado em contadores de *hardware* na AWS. Primeiro são apresentadas as informações sobre a infraestrutura e as aplicações utilizadas nos testes. Então, os resultados obtidos são discutidos.

¹<http://github.com/lapesd/hpcac-toolkit>

²A ferramenta HPC@Cloud utiliza a implementação MVAPICH do MPI, disponível em <http://mvapich.cse.ohio-state.edu>

³<https://github.com/torvalds/linux/tree/master/tools/perf>

3.1. Infraestrutura e Aplicações

Foram utilizadas duas máquinas nos experimentos, ambas equipadas com processadores *Intel Xeon Platinum* e com sistema operacional Amazon Linux 2, porém com configurações ligeiramente diferentes⁴: (i) *c5n.9xlarge*: instância virtualizada da AWS composta por um processador Intel Xeon Platinum 8124M de 3.00GHz com 18 núcleos físicos (36 *threads* com *hyperthreading*) e 96 GB de memória RAM; e (ii) *c5.metal*: instância *baremetal* (não virtualizada) da AWS composta por dois processadores Intel Xeon Platinum 8275CL de 3.00GHz com 24 núcleos físicos cada (total de 96 *threads* com *hyperthreading*) e 192 GB de memória RAM.

As aplicações executadas são provenientes do *NAS Parallel Benchmarks*⁵ [Bailey et al. 1991] (versão 3.4.2) e todos os testes foram executados com a classe de problema “B” e com as versões paralelas das aplicações implementadas com *Message Passing Interface* (MPI). Cada aplicação foi executada com 2, 4, 8, 16 e 32 processos MPI. A ferramenta Perf foi utilizada de forma a monitorar cada processo MPI individualmente. Os resultados apresentados representam a média dos valores coletados em cada um dos processos MPI. Foram realizados experimentos com três aplicações do NAS bastante distintas: *Embarrassingly Parallel* (EP), *Integer Sort* (IS) e *3D Fourier Transform* (FT). A aplicação EP se destaca por quase não haver dependência ou comunicação entre os processos, servindo, portanto, para evidenciar parte do impacto do *hypervisor* em comparação aos casos sem virtualização (*baremetal*). A aplicação IS realiza a ordenação de valores inteiros utilizando o algoritmo *bucket sort* e apresenta um grau moderado de comunicação entre os processos MPI. Por fim, a aplicação FT resolve numericamente equações diferenciais parciais usando transformada rápida de Fourier e é caracterizada por muitas comunicações do tipo ‘todos para todos’. Os testes foram realizados uma vez cada.

As métricas coletadas com a ferramenta Perf foram: *CPU Clock* - que coleta a quantidade de ciclos necessários para cumprir a execução da aplicação; *CPU Task Clock* - que mede em milissegundos o tempo que a CPU demora para realizar a execução da aplicação; *Cache References* - que informa o total de requisições à memória *cache* e *Cache Misses* - que coleta a quantidade de vezes em que uma informação foi solicitada na *cache* mas não foi encontrada (faltas de *cache*).

3.2. Análise dos Resultados

Os resultados preliminares obtidos são apresentados na Figura 1. De forma geral, observa-se que as aplicações apresentam desempenho um pouco superior quando executadas na instância sem virtualização (*c5.metal*). Todavia, não é possível determinar com precisão que esta diferença refere-se somente a um sobrecusto oriundo da virtualização, pois as instâncias *c5.metal* e *c5n.9xlarge* possuem algumas diferenças em suas configurações, apesar de ambas utilizarem processadores Intel de mesma geração.

Com relação às métricas de *cache*, nota-se que, de maneira geral, as aplicações apresentaram um percentual de faltas *cache* bastante similar nos dois tipos de instâncias (*baremetal* e virtualizada), com exceção da aplicação IS que apresentou uma diferença mais significativa.

⁴Não havia disponibilidade de instâncias virtualizadas e não virtualizadas com *hardware* idêntico no momento em que este artigo foi elaborado.

⁵<https://www.nas.nasa.gov/software/npb.html>

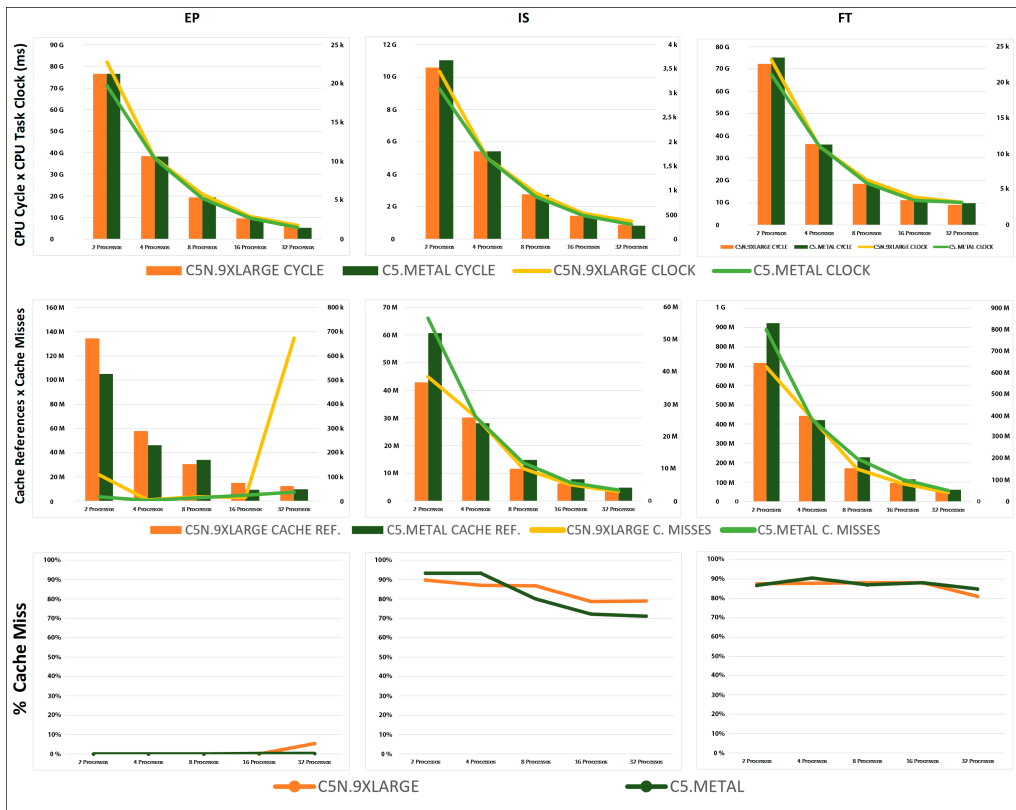


Figura 1. Resultados dos benchmarks EP, IS e FT do NAS.

4. Conclusão

O trabalho mostrou que é possível fazer o perfilamento de aplicações baseado em contadores de *hardware* na AWS, apesar das limitações impostas pelo provedor de nuvem. Os resultados obtidos com as métricas selecionadas mostraram pouca diferença entre as instâncias virtualizada e a não virtualizada utilizadas, com exceção da aplicação IS para a métrica de *cache*. Futuramente, almeja-se realizar experimentos mais completos com outras aplicações do NAS, com instâncias virtualizadas e não virtualizadas da AWS que utilizem exatamente o mesmo *hardware* e realizando diversas repetições.

Referências

- Bailey, D. H., Barszcz, E., Barton, J. T., Browning, D. S., Carter, R. L., Dagum, L., Fatoohi, R. A., Frederickson, P. O., Lasinski, T. A., Schreiber, R. S., et al. (1991). The nas parallel benchmarks. *The International Journal of Supercomputing Applications*, 5(3):63–73.
- Mell, P., Grance, T., et al. (2011). The nist definition of cloud computing.
- Munhoz, V. and Castro, M. (2022). HPC@Cloud: A provider-agnostic software framework for enabling hpc in public cloud platforms. In *Anais do XXIII Simpósio em Sistemas Computacionais de Alto Desempenho (WSCAD)*, pages 157–168. SBC.
- Saini, R. and Sainis, N. (2021). Enable high performance computing in cloud: A review. *International Journal of Scientific Research*, 10:44–45.
- Varghese, B. and Buyya, R. (2018). Next generation cloud computing: New trends and research directions. *Future Generation Computer Systems*, 79:849–861.