

Avaliando o Impacto da Rede no Desempenho e Custo de Execução de Aplicações HPC

Anderson M. Maliszewski^{1,2}, Eduardo Roloff¹, Dalvan Griebler^{2,3},
Philippe O. A. Navaux¹

¹Instituto de Informática, Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre – RS – Brasil

²Laboratório de Pesquisas Avançadas para Computação em Nuvem (LARCC)
Faculdade Três de Maio (SETREM), Três de Maio – RS – Brasil

³Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
Porto Alegre – RS – Brasil

{ammaliszewski, eroloff, navaux}@inf.ufrgs.br

dalvan.griebler@acad.pucrs.br

Resumo. *O desempenho das aplicações HPC depende de dois componentes principais; poder de processamento e interconexão de rede. Este artigo avalia o impacto que a interconexão de rede exerce em programas paralelos usando um cluster homogêneo, em relação a desempenho e custo de execução estimado.*

1. Introdução

Uma demanda crescente por poder computacional é observada devido ao aumento do tamanho de problemas computacionais. Para prover tais recursos, a Computação de Alto Desempenho (HPC) tem sido historicamente usada para acelerar o processamento de dados. Como esses sistemas geralmente usam MPI como padrão para o desenvolvimento de novas aplicações, a interconexão de rede tem uma forte influência no desempenho. Muitos avanços, bem como novos padrões de rede, foram criados para atender a esses requisitos. Entretanto, a interconexão de rede ainda está diretamente relacionada a gargalos no desempenho de aplicações executadas em sistemas distribuídos [Roloff et al. 2017].

Desta forma, foram avaliadas as interconexões InfiniBand FDR (IB FDR) e Gigabit Ethernet (GbE), usando o mesmo *cluster* físico com quatro nós computacionais alocados com a ferramenta Slurm. Cada nó é composto por dois processadores Intel Xeon E5-2650 v3 (Q3'14) Haswell 2,3 GHz, 20 núcleos (10 por CPU) com *Hyper-Threading* ativado, resultando em 40 threads e 128GB de RAM DDR4. Os nós são conectados usando os comutadores Mellanox SX6036 FDR e HPE 1820-48G. Foram selecionadas duas aplicações (EP e FT) do conjunto NPB [Bailey et al. 1991], a aplicação real Alya [Vázquez et al. 2016] e uma aplicação (CPU-8levels) do conjunto Imbalanced Benchmarks [Roloff. et al. 2018]. Os testes foram executados com 30 replicações usando o *cluster* físico sem custos de alocação. No entanto, sabendo que o custo é um dos problemas mais relevantes para a computação em nuvem, o cenário mais próximo do mostrado neste trabalho disponível na nuvem pública foi estimado. O provedor Microsoft Azure foi usado porque possui as instâncias A10 (US\$ 0,78 por hora) e A8 (US\$ 0,975 por hora), com as mesmas especificações de hardware (CPU, Memória) e diferentes interconexões de rede (instância A8 utiliza InfiniBand e a instância A10 usa Gigabit Ethernet).

Aplicação	Tempo de Exec GbE (s)	Tempo de Exec IB (s)	Custo de Exec GbE	Custo de Exec IB
Alya	297,06	254,12	US\$ 0,26	US\$ 0,28
EP	37,39	37,30	US\$ 0,03	US\$ 0,04
FT	1580,88	145,36	US\$ 1,36	US\$ 0,16
IMB-CPU	9,04	8,95	US\$ 0,0078	US\$ 0,0097

Tabela 1. Tempo e custo de execução das aplicações Alya, EP, FT e IMB-CPU usando as interconexões Gigabit Ethernet (GbE) e InfiniBand (IB FDR).

2. Resultados

Os resultados, apresentados na Tabela 1, mostram que a aplicação Alya tem uma melhoria de desempenho e custo de $\approx 16\%$ e $\approx 6,4\%$, respectivamente, usando a interconexão InfiniBand FDR. As aplicações EP e IMB-CPU não apresentam ganhos de desempenho usando o IB FDR e, conseqüentemente, pior resultado no custo de execução, com uma variação de $\approx 20\%$, resultado esperado, uma vez que elas são CPU-Bound. Por outro lado, mesmo considerando que a instância A8 com IB é 25% mais cara, a aplicação FT mostra um enorme ganho de desempenho (até $\approx 989\%$) e custo ($\approx 770\%$) usando o IB FDR. Essa aplicação é dominada pela comunicação, com *ranks* MPI trocando várias pequenas mensagens e, portanto, a latência desempenha um papel crucial no ganho de desempenho.

3. Conclusões

Nos resultados, demonstrou-se que a interconexão desempenha um papel crucial nas aplicações MPI que possuem maior utilização da rede. Por exemplo, a aplicação FT executa $\approx 989\%$ melhor com o InfiniBand FDR em comparação com a interconexão Gigabit Ethernet, o que leva a um custo de execução mais barato em $\approx 770\%$. Por outro lado, para aplicações CPU-Bound, como EP e IMB-CPU, uma interconexão mais rápida pode não ser adequada para ter um desempenho aprimorado, o que também torna sua execução em instâncias com InfiniBand financeiramente inviável. Para o futuro, planeja-se realizar a mesma avaliação nos principais provedores de nuvem pública, considerando as interconexões de rede disponíveis com foco no custo de execução.

Agradecimentos Este trabalho foi parcialmente apoiado pelos projetos FAPERGS GREEN-CLOUD (No 16/2551-0000 488-9) e PARAELASTIC (No. 17/2551-0000871-5).

Referências

- Bailey, D. H., Barszcz, J. T., Barton, J., Browning, D. S., Carter, R. L., Dagum, L., Fatoohi, R. A., Frederickson, P. O., Lasinski, T., Schreiber, R., Simon, H., Venkatakrishnan, V., and Weeratunga, S. (1991). The NAS Parallel Benchmarks; Summary and Preliminary Results. In *ACM/IEEE Conference on Supercomputing (SC)*.
- Roloff, E., Diener, M., Gaspary, L. P., and Navaux, P. O. A. (2017). HPC Application Performance and Cost Efficiency in the Cloud. In *Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*.
- Roloff, E., Diener, M., Gaspary, L. P., and Navaux, P. O. A. (2018). Exploiting Load Imbalance Patterns for Heterogeneous Cloud Computing Platforms. In *International Conference on Cloud Computing and Services Science (CLOSER)*.
- Vázquez, M., Houzeaux, G., Koric, S., Artigues, A., Aguado-Sierra, J., Arís, R., Mira, D., Calmet, H., Cucchiatti, F., Owen, H., et al. (2016). Alya: Multiphysics Engineering Simulation Toward Exascale. *Journal of Computational Science*.