

Otimizando a Localização de MVs na Nuvem Através de Padrões de Comunicação

Artur Baruchi

October 23, 2015

1 Abstract

O uso de computação na nuvem por usuários de Computação de Alto Desempenho (HPC - *High Performance Computing*) ainda sofre muita resistência, em grande medida, por conta da grande degradação de desempenho que essas aplicações sofrem ao serem executadas na nuvem. Entre os principais fatores que levam à degradação, pode-se citar (1) interconexão de rede lenta, que utiliza hardware commodity, (2) ambiente heterôgeneo com diferentes gerações de processadores, (3) compartilhamento do hardware entre diferentes usuários e (4) sobrecarga da virtualização. As estratégias empregadas com objetivo de amenizar as limitações da nuvem estão situadas em duas grandes linhas de pesquisa, a primeira é tornar as aplicações conscientes de estarem sendo executadas em uma ambiente de nuvem (HPC-Aware) e a segunda é tornar a nuvem ciente de que uma determinada aplicação é HPC e, dessa forma, se adapta para melhor atender os requisitos (Cloud-Aware).

Este trabalho situa-se na segunda linha de pesquisa (Cloud-Aware) e tem como objetivo detectar padrões da comunicação de aplicações HPC e com isso alocar as Máquinas Virtuais (VM - Virtual Machine) de maneira a reduzir ou eliminar a sobrecarga inerente de um ambiente de computação. Ao reconhecer um determinado padrão de comunicação poderia ser beneficiado por uma proximidade física das MVs (no mesmo hospedeiro físico ou no mesmo rack) o orquestrador faria a migração da(s) MV(s) com o objetivo de aproximá-las. Outra estratégia seria determinar quais padrões de comunicação são conflitantes e, com isso, evitar que MVs sejam alocadas no mesmo hospedeiro.

2 Keywords

HPC, Cloud, Communication, Message Passing, Characterization

3 Introdução

A adesão de usuários de aplicações de Computação de Alto Desempenho (HPC - *High Performance Computing*), seja no âmbito acadêmico ou corporativo não é comum. Isso ocorre devido a diversos fatores, como interconexão de rede lenta, ambiente computacional heterogêneo, sobrecarga da virtualização e compartilhamento de recursos. De fato, a computação na nuvem tem como principal argumento a redução de custos de manter um parque computacional complexo e, de forma geral, centros computacionais que tem a sua disposição supercomputadores são menos suscetíveis ao argumento do custo. Entretanto, usuários de aplicações HPC em centros computacionais de pequeno e médio porte (ex. Startups, pequenos laboratórios de pesquisa, etc) possuem maior sensibilidade a apelo da computação da nuvem e por essa razão mais suscetíveis a adotá-la.

Apesar do custo da computação na nuvem ser um atrativo, sua adoção ainda sofre grande resistência por parte de usuários de pequeno e médio porte de aplicações HPC em razão da degradação do desempenho. Em grande medida, a degradação de desempenho de aplicações HPC na nuvem deve-se ao uso de compartilhado de recursos computacionais, principalmente de interfaces de rede.

- 4 Revisão
- 5 Conclusão
- 6 Acknowledgments
- 7 Abbreviations
- 8 Author Information
- 9 Author contributions
- 10 Competing Interests
- 11 Cover Letter