

Monitoramento Preditivo de Recursos em Cloud Computing

Júlio Cezar Santos Pires, Cristiano André da Costa, Rodrigo da Rosa Righi

¹Programa de Pós-Graduação em Computação Aplicada - PIPCA

Universidade do Vale do Rio dos Sinos - UNISINOS

Av. Unisinos, 950 - Bairro Cristo Rei - CEP 93.022-000 - São Leopoldo – RS – Brasil

juliocspires@gmail.com, {cac, rrrighi}@unisinos.br

***Resumo.** Este trabalho aborda a área de monitoramento de recursos em Computação em Nuvem, utilizando para tal a coleta e análise de padrões de desempenho dos recursos nos três níveis da infraestrutura da Nuvem (máquina física, máquina virtual e processo). Neste sentido tem-se o monitoramento hierárquico, visando melhores resultados para tomada de decisão em situações de migração, provisionamento e consolidação dos recursos.*

1. Introdução

Segundo [Mell and Grance 2011], *Cloud Computing* (ou Computação em Nuvem) é um modelo para permitir o acesso à rede sob demanda a um conjunto compartilhado de recursos computacionais configuráveis, tais como: redes, servidores, armazenamento, aplicativos e serviços, que podem ser rapidamente provisionados e liberados com o mínimo esforço de gerenciamento ou interação com o provedor de serviços.

Entre as principais características da Computação em Nuvem, tem-se a elasticidade, provisionamento de serviço, cobrança baseada na utilização efetiva dos recursos, entre outras. Em busca de tornar estas características na prática possíveis, visando um controle mais fino sobre os recursos computacionais, torna-se necessário que a infraestrutura da Nuvem disponha de um sistema de monitoramento que efetue a constante inspeção sob os recursos disponíveis.

Este trabalho aborda maneiras de oferecer predição de desempenho de recursos em Computação em Nuvem. Nesse sentido, tem-se o monitoramento preditivo hierárquico, ou seja, atuando nos três níveis da infraestrutura da Nuvem (máquina física, máquina virtual e aplicação (processo)) de forma preditiva, desta forma, utilizando os padrões de utilização dos recursos como objeto de pesquisa para se atingir esse objetivo.

2. Motivação e Passos Iniciais da Pesquisa

Com o uso de sistemas de monitoramento de recursos em *Cloud Computing*, torna-se possível detectar oportunidades de ganho, seja este em relação ao aumento do desempenho através do provisionamento e migração de recursos ou relacionado à receita, neste caso economizando energia por meio da consolidação de recursos na infraestrutura da Nuvem. Além disso, o sistema de monitoramento detectará falhas de SLA, conforme as coletas e análises periódicas dos recursos, visando à verificação e o cumprimento do contrato firmado entre cliente e prestador de serviço.

Desta forma, com um sistema monitoramento preditivo atuando nos três níveis da infraestrutura da Nuvem, torna-se possível a identificação e predição de processos com

baixo desempenho ou situações muito discrepantes com a realidade, visto que o mesmo tende a ser a ser mais eficiente devido à combinação dos três níveis da hierarquia dos recursos computacionais presente na infraestrutura da Nuvem.

Além de preditivo e hierárquico, o sistema de monitoramento de recursos deve ser também proativo e assim, estar sempre alinhado com as ações que podem ser tomadas através da combinação de dados de monitoramento de uma ou mais fontes de informações de desempenho. De posse de boas métricas de monitoramento, tem-se meios para a tomada de decisão e assim efetuar a migração de processos e máquinas virtuais, ganhar desempenho, consolidar recursos, economizar energia e assim, reduzir custos.

Como gerenciador da infraestrutura, será utilizado o *middleware* de Computação em Nuvem OpenNebula. As aplicações a serem executadas dentro deste ambiente serão aplicações de alto desempenho do tipo BSP (*Bulk Synchronous Parallel*), que são aplicações paralelas onde possuem N processos com uma série de super-etapas (computação, comunicação e sincronização). Um dos benefícios da análise e predição de padrões de uso neste caso, é que através da descoberta dos padrões de comunicação e computação, tem-se meios para decidir qual será o processo e/ou máquina virtual candidato(a) para migração.

Desta forma, em um ambiente dinâmico, como uma infraestrutura de Computação em Nuvem, torna-se necessário a análise histórica das métricas coletadas, visto que um dado isolado por si só não é capaz de fazer o sistema efetuar a melhor tomada de decisão no momento do provisionamento e/ou consolidação dos recursos. Esta etapa de análise de padrões envolve o uso de séries temporais. Uma série temporal é uma sequência de números reais, onde cada número representa um valor de dados em um ponto no tempo[Wu et al. 2000]. Para a análise sobre a sequência dos pontos, podem ser utilizadas técnicas de *Pattern Matching*, tais como DFT (*Discrete Fourier Transform*) e DTW (*Dynamic Time Warping*).

3. Conclusões e Trabalhos Futuros

Ações baseadas na predição de recursos, além do controle sobre a disponibilidade dos recursos podem ser úteis nas seguintes situações: (i) projeção de desempenho futuro, (ii) tratamento da energia consumida pela Nuvem; (iii) migração de máquinas virtuais; (iv) migração de processos; (v) auxílio na administração dos recursos; (vi) possibilidade de estimar o tempo para executar um serviço e; (vii) gerenciamento do SLA entre consumidores e o provedor de serviço. Neste sentido, trabalhos futuros compreendem um estudo mais aprofundado de técnicas para análise de séries temporais, análise nas taxas de acerto da função de predição, além de métodos computacionais para a interpretação e processamento de padrões de aplicações paralelas do tipo BSP.

Referências

- Mell, P. and Grance, T. (2011). The nist definition of cloud computing. *NIST special publication*, 800:145.
- Wu, Y.-L., Agrawal, D., and El Abbadi, A. (2000). A comparison of dft and dwt based similarity search in time-series databases. In *Proceedings of the ninth international conference on Information and knowledge management, CIKM '00*, pages 488–495, New York, NY, USA. ACM.