

Aprimorando a elasticidade em ambientes de virtualização em nível de Sistema Operacional

Israel Campos de Oliveira

Orientador : Prof. Dr. César Augusto F. De Rose

Introdução

- Computação em Nuvem e Elasticidade:
Infraestrutura com capacidade de provisionamento de recursos de forma rápida e **elástica**, em alguns casos **automática**, para **aumentar e diminuir recursos**. [1]
- Estratégias de elasticidade:
 - Replicação
 - Redimensionamento

Desafios - Elasticidade

- Provedores como Amazon AWS, oferecem elasticidade como serviço, porém:
 - Dependem da ação do usuário para definir regras, condições (limites) e ações
 - Complexidade para a criação de regras
 - Regras criadas através de limites que apresentam valores fixos permitem a quebra de SLAs quando o **comportamento da aplicação se mostra irregular**

Oportunidades

- Elasticidade em nível de Sistema Operacional:
 - Menor overhead
- Aplicações de banco de dados
 - Tendência de migração para a nuvem.

Objetivos

Avaliar o impacto do uso da virtualização em nível de Sistema Operacional sobre a estratégia de alocação dinâmica de recursos proposta por Dawoud et al. [11], com foco no uso de aplicações de banco de dados e com a finalidade de manter um determinado SLA.

Avaliações Arquitetura

Somente Leitura

SLA (TPS)	Total de Transações	Média de TPS	SLA Atendido	Quebra SLA
1300	429464	1431,54	99,4%	0.6%

Somente Leitura / Escrita

SLA (TPS)	Total de Transações	Média de TPS	SLA Atendido	Quebra SLA
1200	387369	1291,23	87%	13%

Conclusão

Os resultados obtidos são promissores, pois mostrou-se que é possível utilizar uma técnica reativa para gerir o processo de elasticidade, minimizando situações de thrashing .

Os resultados mostram que no melhor caso a manutenção do SLO ocorre em mais de 99% do tempo de execução das aplicações e no pior caso houve uma quebra de SLOs em cerca de 13%, devido a limitações dessa proposta quanto a gerência de E/S

Perguntas ?



Referências

- [1] T. G. Mell Peter, “The nist definition of cloud computing,” Special Publication 800-145, 2011
- [2] B. Suleiman, S. Sakr, R. Jeffery, and A. Liu, “On understanding the economics and elasticity challenges of deploying business applications on public cloud infrastructure,” *Journal of Internet Services and Applications*, vol. 3, no. 2, pp. 173–193, 2012. [Online]. Available: <http://dx.doi.org/10.1007/s13174-011-0050-y>
- [3] T. Chieu, A. Mohindra, A. Karve, and A. Segal, “Dynamic scaling of web applications in a virtualized cloud computing environment,” in *e-Business Engineering*, 2009. ICEBE '09. IEEE International Conference on, Oct 2009, pp. 281–286.
- [4] X. Zhu, Z. Wang, and S. Singhal, “Utility-driven workload management using nested control design,” in *American Control Conference*, 2006, June 2006, pp. 6 pp.—.
- [5] M. Xavier, M. Neves, F. Rossi, T. Ferreto, T. Lange, and C. De Rose, “Performance evaluation of container-based virtualization for high performance computing environments,” in *Parallel, Distributed and Network-Based Processing (PDP)*, 2013 21st Euromicro International Conference on, Feb 2013, pp. 233–240.
- [6] T. Lange, P. Cemim, M. Xavier, and C. De Rose, “Optimizing the management of a database in a virtual environment,” in *Computers and Communications (ISCC)*, 2013 IEEE Symposium on, July 2013, pp. 000 594–000 599.