

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Interference Aware Scheduling for Cloud Computing

Diogo Trindade Basto

VERSÃO DE TRABALHO



Mestrado Integrado em Engenharia Informática e Computação

Orientador: Jorge Manuel Gomes Barbosa

11 de Fevereiro de 2014

Interference Aware Scheduling for Cloud Computing

Diogo Trindade Basto

Mestrado Integrado em Engenharia Informática e Computação

Resumo

O Resumo fornece ao leitor um sumário do conteúdo da dissertação. Deverá ser breve mas conter detalhe suficiente e, uma vez que é a porta de entrada para a dissertação, deverá dar ao leitor uma boa impressão inicial.

Este texto inicial da dissertação é escrito no fim e resume numa página, sem referências externas, o tema e o contexto do trabalho, a motivação e os objectivos, as metodologias e técnicas empregues, os principais resultados alcançados e as conclusões.

Este documento ilustra o formato a usar em dissertações na Faculdade de Engenharia da Universidade do Porto. São dados exemplos de margens, cabeçalhos, títulos, paginação, estilos de índices, etc. São ainda dados exemplos de formatação de citações, figuras e tabelas, equações, referências cruzadas, lista de referências e índices. É usado texto descartável, *Loren Ipsum*, para preencher a dissertação por forma a ilustrar os formatos.

Seguem-se umas notas breves mas muito importantes sobre a versão provisória e a versão final do documento. A versão provisória, depois de verificada pelo orientador e de corrigida em contexto pelo autor, deve ser publicada na página pessoal de cada estudante/dissertação, juntamente com os dois resumos, em português e em inglês; deve manter a marca da água, assim como a numeração de linhas conforme aqui se demonstra.

A versão definitiva, a produzir somente após a defesa, em versão impressa (dois exemplares com capas próprias FEUP) e em versão eletrónica (6 CDs com "rodela" própria FEUP), deve ser limpa da marca de água e da numeração de linhas e deve conter a identificação, na primeira página, dos elementos do júri respetivo. Deve ainda, se for o caso, ser corrigida de acordo com as instruções recebidas dos elementos júri.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed vehicula lorem commodo dui. Fusce mollis feugiat elit. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec eu quam. Aenean consectetur odio quis nisi. Fusce molestie metus sed neque. Praesent nulla. Donec quis urna. Pellentesque hendrerit vulputate nunc. Donec id eros et leo ullamcorper placerat. Curabitur aliquam tellus et diam.

Ut tortor. Morbi eget elit. Maecenas nec risus. Sed ultricies. Sed scelerisque libero faucibus sem. Nullam molestie leo quis tellus. Donec ipsum. Nulla lobortis purus pharetra turpis. Nulla laoreet, arcu nec hendrerit vulputate, tortor elit eleifend turpis, et aliquam leo metus in dolor. Praesent sed nulla. Mauris ac augue. Cras ac orci. Etiam sed urna eget nulla sodales venenatis. Donec faucibus ante eget dui. Nam magna. Suspendisse sollicitudin est et mi.

Phasellus ullamcorper justo id risus. Nunc in leo. Mauris auctor lectus vitae est lacinia egestas. Nulla faucibus erat sit amet lectus varius semper. Praesent ultrices vehicula orci. Nam at metus. Aenean eget lorem nec purus feugiat molestie. Phasellus fringilla nulla ac risus. Aliquam elementum aliquam velit. Aenean nunc odio, lobortis id, dictum et, rutrum ac, ipsum.

Ut tortor. Morbi eget elit. Maecenas nec risus. Sed ultricies. Sed scelerisque libero faucibus sem. Nullam molestie leo quis tellus. Donec ipsum.

Abstract

Here goes the abstract written in English.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed vehicula lorem commodo dui. Fusce mollis feugiat elit. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec eu quam. Aenean consectetur odio quis nisi. Fusce molestie metus sed neque. Praesent nulla. Donec quis urna. Pellentesque hendrerit vulputate nunc. Donec id eros et leo ullamcorper placerat. Curabitur aliquam tellus et diam.

Ut tortor. Morbi eget elit. Maecenas nec risus. Sed ultricies. Sed scelerisque libero faucibus sem. Nullam molestie leo quis tellus. Donec ipsum. Nulla lobortis purus pharetra turpis. Nulla laoreet, arcu nec hendrerit vulputate, tortor elit eleifend turpis, et aliquam leo metus in dolor. Praesent sed nulla. Mauris ac augue. Cras ac orci. Etiam sed urna eget nulla sodales venenatis. Donec faucibus ante eget dui. Nam magna. Suspendisse sollicitudin est et mi.

Fusce sed ipsum vel velit imperdiet dictum. Sed nisi purus, dapibus ut, iaculis ac, placerat id, purus. Integer aliquet elementum libero. Phasellus facilisis leo eget elit. Nullam nisi magna, ornare at, aliquet et, porta id, odio. Sed volutpat tellus consectetur ligula. Phasellus turpis augue, malesuada et, placerat fringilla, ornare nec, eros. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Vivamus ornare quam nec sem mattis vulputate. Nullam porta, diam nec porta mollis, orci leo condimentum sapien, quis venenatis mi dolor a metus. Nullam mollis. Aenean metus massa, pellentesque sit amet, sagittis eget, tincidunt in, arcu. Vestibulum porta laoreet tortor. Nullam mollis elit nec justo. In nulla ligula, pellentesque sit amet, consequat sed, faucibus id, velit. Fusce purus. Quisque sagittis urna at quam. Ut eu lacus. Maecenas tortor nibh, ultricies nec, vestibulum varius, egestas id, sapien.

Phasellus ullamcorper justo id risus. Nunc in leo. Mauris auctor lectus vitae est lacinia egestas. Nulla faucibus erat sit amet lectus varius semper. Praesent ultrices vehicula orci. Nam at metus. Aenean eget lorem nec purus feugiat molestie. Phasellus fringilla nulla ac risus. Aliquam elementum aliquam velit. Aenean nunc odio, lobortis id, dictum et, rutrum ac, ipsum.

Ut tortor. Morbi eget elit. Maecenas nec risus. Sed ultricies. Sed scelerisque libero faucibus sem. Nullam molestie leo quis tellus. Donec ipsum. Nulla lobortis purus pharetra turpis. Nulla laoreet, arcu nec hendrerit vulputate, tortor elit eleifend turpis, et aliquam leo metus in dolor. Praesent sed nulla. Mauris ac augue. Cras ac orci. Etiam sed urna eget nulla sodales venenatis. Donec faucibus ante eget dui. Nam magna. Suspendisse sollicitudin est et mi.

Phasellus ullamcorper justo id risus. Nunc in leo. Mauris auctor lectus vitae est lacinia egestas. Nulla faucibus erat sit amet lectus varius semper. Praesent ultrices vehicula orci. Nam at metus. Aenean eget lorem nec purus feugiat molestie. Phasellus fringilla nulla ac risus. Aliquam elementum aliquam velit. Aenean nunc odio, lobortis id, dictum et, rutrum ac, ipsum.

Ut tortor. Morbi eget elit. Maecenas nec risus. Sed ultricies. Sed scelerisque libero faucibus sem. Nullam molestie leo quis tellus. Donec ipsum.

Conteúdo

1	Introdução	1
1.1	Motivação e Objetivos	1
1.2	Estrutura da Dissertação	2
2	Estado da Arte	3
2.1	Computação em Nuvem	3
2.2	Escalonamento de Tarefas em Nuvens de Computação	3
2.3	Tecnologias de Virtualização	4
2.3.1	Interferências em Tecnologias de Virtualização	4
2.4	SimGrid	5
2.5	Conclusões	5
	Referências	7

CONTEÚDO

Lista de Figuras

LISTA DE FIGURAS

Lista de Tabelas

LISTA DE TABELAS

Abreviaturas e Símbolos

API	Application Programming Interface
CERN	Conseil Européen pour la Recherche Nucléaire
KVM	Kernel-based Virtual Machine
NSA	National Security Agency
QoS	Quality of Service
SLA	Service Level Agreement

Capítulo 1

Introdução

Nuvens computacionais são um serviço emergente na indústria tecnológica, pois permitem computação a um custo reduzido. A própria tecnologia é apenas a adaptação de outras tecnologias existentes para responder a um crescimento de utilização de serviços na *internet*. Este serviço disponibilizado ao público por empresas como a Google, Amazon ou Microsoft permite a expansão do poder computacional aos utilizadores conforme as suas necessidades, eliminando um investimento inicial e possibilitando o crescimento do serviço. Grandes organizações, como o caso do CERN e NSA, utilizam redes computacionais internas para disponibilizar o poder computacional que as suas tarefas diárias necessitam, não limitando cada colaborador ao poder do seu computador pessoal, permitindo a aproveitação dos recursos organizacionais para a conclusão de tarefas mais rapidamente.

A presente dissertação apresentará um método para melhorar o serviço das nuvens computacionais através dum escalonamento de tarefas mais consciente sobre o processo de virtualização e das interferências que o mesmo provoca na realização de tarefas nos nós computacionais.

1.1 Motivação e Objetivos

A otimização da utilização dos recursos disponíveis permite melhorias na qualidade do serviço (QoS) fornecido aos utilizadores, através de serviços mais rápidos e responsivos, e menor gastos por parte dos fornecedores, através de reduções energéticas devido a menores períodos de utilização e cumprimento dos acordos da qualidade do serviço (SLA), influenciando positivamente todas as partes interessadas. É, portanto, uma área de interesse estudar otimizações possíveis que contribuam para a eficiência do serviço. Uma parte principal do funcionamento de uma nuvem é o escalonador, responsável por atribuir as tarefas aos nós disponíveis. Esta atribuição é feita em linha, aquando da chegada de uma tarefa, o escalonador deverá avaliar os vários nós da rede computacional e despachar a tarefa para o nó que possua a melhor avaliação. Esta avaliação tem em conta as características de uma tarefa, através da especificação do pedido ou por execução da

mesma temporariamente e medir os recursos exigidos, e as capacidades e quantidade de trabalho de cada nó. Técnicas de escalonamento atuais não tem em conta a interação entre tarefas, um dos principais culpados nos atrasos na nuvem, devido às tecnologias de virtualização e arquitetura do *hardware* usados. Quando duas tarefas com utilizações de recursos semelhantes, por exemplo ambas tem bastante acesso a disco para recolha e escrita de dados com pouco tratamento dos mesmos a nível do processador, são atribuídas ao mesmo nó, que terá espaço de disco para ambas, as máquinas virtuais irão interferir entre si no acesso a recursos que não podem ser divididos para cada virtualização devido à sua arquitetura, neste caso, o barramento de dados.

Esta dissertação tem como objetivo desenvolver um escalonador que utilize métricas de interferência para poder reduzir o impacto das mesmas na nuvem. Para isso será estudada a interferência entre vários tipos de tarefas, em bibliografia existente, e utilizado um simulador de uma nuvem computacional para avaliar o desempenho do escalonador. Para efeitos de avaliação de resultados será utilizado um escalonador de código aberto, utilizado atualmente na indústria, o *Filter Scheduler* do OpenStack [Gon12]. Para simulação será utilizado o SimGrid e a arquitetura de rede será idêntica à utilizada no CERN, podendo para simulação de tarefas ser utilizado os dados de utilização da sua nuvem ou criadas novas tarefas de uso científico geral, como especificadas em [Meh13].

1.2 Estrutura da Dissertação

Os capítulos seguintes deste relatório encontram-se da forma seguinte. O capítulo 2 apresenta o estado da arte na área de computação em nuvem, sendo específico na área de escalonamento e de interferências na virtualização. O capítulo 3 apresenta os dados sobre interferência recolhidos e explica o impacto dos mesmos na rede computacional. O capítulo 4 descreve detalhes de implementação que serão utilizados para a realização do trabalho e os métodos utilizados para avaliação de resultados. O capítulo 5 apresentará as conclusões retiradas da preparação de dissertação e o plano de trabalho a seguir.

Capítulo 2

Estado da Arte

O presente capítulo identificará referências bibliográficas sobre o presente estado da computação em nuvem, as suas utilizações e proliferação. De seguida, apresenta trabalhos anteriores sobre escalonamento de tarefas na nuvem, seguido de análise de trabalhos sobre interferências entre máquinas virtuais e sobre o simulador a utilizar, SimGrid.

2.1 Computação em Nuvem

No artigo [AFG⁺10], os autores apresentam bastante sucintamente a utilidade dos serviços de computação oferecidos atualmente e os principais obstáculos e oportunidades de desenvolvimento. Em [BYV⁺09] os autores apresentam a nuvem como a futura quinta utilidade e o modelo de negócio que é utilizado pelos fornecedores de serviço ainda hoje. Para o cumprimento dos contratos estipulados e combater os atrasos são provisionados recursos a mais dos que foram contratados, diminuindo a eficiência da nuvem, como explicado em [AFG⁺10], [NKG10] e [CFF12]. Em [WGL⁺12] é feita a comparação entre os dois principais programas de gestão de nuvens, OpenStack e OpenNebula. Os autores de [VPB09] demonstram a utilidade da computação em nuvem na comunidade científica, apresentando um estudo sobre a utilização do poder computacional na classificação de dados de expressão genética e compilação de imagens de ressonâncias magnéticas. Devido à disponibilização dos dados e configuração do CERN para a realização desta dissertação, foi escolhido o OpenStack por ser o mesmo utilizado pela organização [Fou13]. O algoritmo a ser replicado para efeitos de teste e avaliação de resultados está descrito em [Gon12].

2.2 Escalonamento de Tarefas em Nuvens de Computação

Existe na área das nuvens de computação bastante bibliografia teórica e prática sobre escalonamento. Em [ABN00] é estudada a utilização de algoritmos probabilísticos, como algoritmos

genéticos, arrefecimento simulado e pesquisa tabu. Em [DK08] é descrito um algoritmo de escalonamento por listagem. No entanto as duas soluções requerem um conhecimento à priori da totalidade das tarefas que vão ser executadas, o que não é possível na maioria dos casos. Em [THW02] os autores apresentam os algoritmos de primeiro tempo de chegada heterogêneo e caminho crítico no processador, algoritmos que são focados na rapidez da resposta. Vários trabalhos focam-se no escalonamento de tarefas de área científica, cuja topologia é conhecida e possa ser dividida entre várias sub-tarefas diferentes. Este escalonamento é NP-completo. Os autores de [GMK12] utilizam uma heurística que aproxima as sub-tarefas dentro da rede, de maneira a acelerar o processo de partilha de dados entre as tarefas e tenta que tarefas iguais fiquem em nós de igual capacidade, para que possam todos trabalhar à sua velocidade máxima. O trabalho em [CZ09] segue a mesma lógica, utilizando um algoritmo de colônia de formigas. Os autores de [SLH⁺13] utilizam ótimos de Pareto para determinar o melhor nó para alocar as tarefas, melhorando sucessivamente a sua solução. Em [ZS13] é demonstrada uma extensão ao algoritmo de escalonamento primeiro tempo de chegada heterogêneo.

2.3 Tecnologias de Virtualização

Virtualização é a tecnologia que permite a um único computador isolar as várias tarefas que são executadas simultaneamente, sendo controlados por um monitor de máquinas virtuais em cada nó, o hipervisor. Em [GZ12] são apresentadas os hipervisores mais utilizados no mercado, programas que criam automaticamente em cada computador as máquinas virtuais necessárias para a execução de tarefas, concluindo com a listagem de problemas comuns entre eles. São também comparados e testados vários hipervisores em [YHB⁺11], obtendo resultados que favorecem o KVM para nuvens de computação de alto rendimento. Em [CGFC11] é sugerido um escalonamento de tarefas a nível de nós individuais, ficando cada nó responsável por garantir a qualidade do serviço, realocando recursos a cada máquina virtual conforme necessário.

2.3.1 Interferências em Tecnologias de Virtualização

A interferência entre máquinas virtuais no mesmo anfitrião começou por ser estudada em [KKB⁺07], apresentando motivos para os atrasos observados, executando testes experimentais para conjuntos de duas máquinas virtuais simultâneas e usando regressão linear para prever conjuntos não testados. O estudo de interferência em tarefas que requerem comunicações dentro da rede em [PLM⁺10], utilizando tarefas limitadas por poder de computacional ou velocidade da rede, demonstra o efeito que a concorrência em componentes partilhados provoca no atraso das tarefas e conclui que a combinação de tarefas com usos de componentes diferentes leva ao menor efeito de interferência. Os autores de [NKG10] determinam que para as qualidade de serviço ser obedecida, cada nó deverá ter um espaço livre para poder alocar mais espaço a uma das virtualizações, não sendo eficiente na utilização dos recursos a troco de oferecer um serviço mais fidedigno. Em [ZT12] são construídos modelos de interferência entre máquinas virtuais, utilizando o modelo e um conjunto de regras para determinar se duas tarefas podem ser consolidadas no mesmo nó.

2.4 SimGrid

SimGrid é a plataforma de simulação a ser utilizada para teste do algoritmo desenvolvido. Os autores de [Cas01] e [LMC03] demonstram casos de uso do simulador para testar algoritmos de escalonamento em versões anteriores da plataforma, podendo-se encontrar em [CLQ08] um estudo com a versão mais recente.

2.5 Conclusões

O estudo da bibliografia revela a importância das nuvens computacionais para a indústria na atualidade. Sendo uma tecnologia ainda em desenvolvimento, apresenta vulnerabilidades já identificadas que podem ser corrigidas para melhorar o serviço atual. Os estudos sobre interferência mostram uma redução de desempenho para valores até um terço do original entre um par de tarefas, podendo-se concluir que as tarefas executadas em série em vez de paralelizadas levariam a um desempenho superior. Esta interferência devido a recursos compartilhados, por exemplo memória *cache*, impossíveis de serem divididos por um hipervisor pode ser combatida com a adaptação de *hardware* que não utilize partilha de recursos ou a adaptação da atribuição de tarefas para verificar a consolidação de tarefas com pouca interferência.

Estado da Arte

Referências

- [ABN00] Ajith Abraham, Rajkumar Buyya e Baikunth Nath. Nature’s heuristics for scheduling jobs on computational grids. In *The 8th IEEE international conference on advanced computing and communications (ADCOM 2000)*, pages 45–52, 2000.
- [AFG⁺10] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica et al. A view of cloud computing. *Communications of the ACM*, 53(4):50–58, 2010.
- [BYV⁺09] Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal, James Broberg e Ivona Brandic. Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation computer systems*, 25(6):599–616, 2009.
- [Cas01] Henri Casanova. Simgrid: A toolkit for the simulation of application scheduling. In *Cluster Computing and the Grid, 2001. Proceedings. First IEEE/ACM International Symposium on*, pages 430–437. IEEE, 2001.
- [CFF12] Antonio Corradi, Mario Fanelli e Luca Foschini. Vm consolidation: a real case based on openstack cloud. *Future Generation Computer Systems*, 2012.
- [CGFC11] Tommaso Cucinotta, Dhaval Giani, Dario Faggioli e Fabio Checconi. Providing performance guarantees to virtual machines using real-time scheduling. In *Euro-Par 2010 Parallel Processing Workshops*, pages 657–664. Springer, 2011.
- [CLQ08] Henri Casanova, Arnaud Legrand e Martin Quinson. Simgrid: A generic framework for large-scale distributed experiments. In *Computer Modeling and Simulation, 2008. UKSIM 2008. Tenth International Conference on*, pages 126–131. IEEE, 2008.
- [CZ09] Wei-Neng Chen e Jun Zhang. An ant colony optimization approach to a grid workflow scheduling problem with various qos requirements. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 39(1):29–43, 2009.
- [DK08] Mohammad I Daoud e Nawwaf Kharma. A high performance algorithm for static task scheduling in heterogeneous distributed computing systems. *Journal of Parallel and distributed computing*, 68(4):399–409, 2008.
- [Fou13] OpenStack Foundation. Cern uses openstack, 2013. URL: <http://www.openstack.org/user-stories/cern/>.
- [GMK12] Abhishek Gupta, Dejan Milojevic e Laxmikant V Kalé. Optimizing vm placement for hpc in the cloud. In *Proceedings of the 2012 workshop on Cloud services, federation, and the 8th open cirrus summit*, pages 1–6. ACM, 2012.

REFERÊNCIAS

- [Gon12] Yong Sheng Gong. Openstack nova-scheduler and its algorithm, April 2012. URL: https://www.ibm.com/developerworks/community/blogs/e93514d3-c4f0-4aa0-8844-497f370090f5/entry/openstack_nova_scheduler_and_its_algorithm27?lang=en.
- [GZ12] Zonghua Gu e Qingling Zhao. A state-of-the-art survey on real-time issues in embedded systems virtualization. *Journal of Software Engineering and Applications*, 5(4):277–290, 2012.
- [KKB⁺07] Younggyun Koh, Rob Knauerhase, Paul Brett, Mic Bowman, Zhihua Wen e Calton Pu. An analysis of performance interference effects in virtual environments. In *Performance Analysis of Systems & Software, 2007. ISPASS 2007. IEEE International Symposium on*, pages 200–209. IEEE, 2007.
- [LMC03] Arnaud Legrand, Loris Marchal e Henri Casanova. Scheduling distributed applications: the simgrid simulation framework. In *Cluster Computing and the Grid, 2003. Proceedings. CCGrid 2003. 3rd IEEE/ACM International Symposium on*, pages 138–145. IEEE, 2003.
- [Meh13] Gaurang Mehta. Workflowgenerator pegasus, October 2013. URL: <https://confluence.pegasus.isi.edu/display/pegasus/WorkflowGenerator>.
- [NKG10] Ripal Nathuji, Aman Kansal e Alireza Ghaffarkhah. Q-clouds: managing performance interference effects for qos-aware clouds. In *Proceedings of the 5th European conference on Computer systems*, pages 237–250. ACM, 2010.
- [PLM⁺10] Xing Pu, Ling Liu, Yiduo Mei, Sankaran Sivathanu, Younggyun Koh e Calton Pu. Understanding performance interference of i/o workload in virtualized cloud environments. In *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*, pages 51–58. IEEE, 2010.
- [SLH⁺13] Sen Su, Jian Li, Qingjia Huang, Xiao Huang, Kai Shuang e Jie Wang. Cost-efficient task scheduling for executing large programs in the cloud. *Parallel Computing*, 2013.
- [THW02] Haluk Topcuoglu, Salim Hariri e Min-you Wu. Performance-effective and low-complexity task scheduling for heterogeneous computing. *Parallel and Distributed Systems, IEEE Transactions on*, 13(3):260–274, 2002.
- [VPB09] Christian Vecchiola, Suraj Pandey e Rajkumar Buyya. High-performance cloud computing: A view of scientific applications. In *Pervasive Systems, Algorithms, and Networks (ISPAN), 2009 10th International Symposium on*, pages 4–16. IEEE, 2009.
- [WGL⁺12] Xiaolong Wen, Genqiang Gu, Qingchun Li, Yun Gao e Xuejie Zhang. Comparison of open-source cloud management platforms: Openstack and opennebula. In *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*, pages 2457–2461. IEEE, 2012.
- [YHB⁺11] Andrew J Younge, Robert Henschel, James T Brown, Gregor von Laszewski, Judy Qiu e Geoffrey C Fox. Analysis of virtualization technologies for high performance computing environments. In *Cloud Computing (CLOUD), 2011 IEEE International Conference on*, pages 9–16. IEEE, 2011.

REFERÊNCIAS

- [ZS13] Wei Zheng e Rizos Sakellariou. Budget-deadline constrained workflow planning for admission control. *Journal of Grid Computing*, pages 1–19, 2013.
- [ZT12] Qian Zhu e Teresa Tung. A performance interference model for managing consolidated workloads in qos-aware clouds. In *Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on*, pages 170–179. IEEE, 2012.