FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Processamento concorrente de Vídeo no MEO Kanal

João Carlos Santos

Mestrado Integrado em Engenharia Informática e Computação

Orientador: Jorge Barbosa

27 de Junho de 2016

Abstract

Nowadays there is a wide variety of solutions for storage and reproduction of video content in the cloud, such as Youtube, Dailymotion, Vimeo and MEO Kanal. The computational effort associated with the storage, classification, decoding and reproduction of these contents is considerable, whether in terms of computational effort in processing these videos or in terms of data transfer operations. When taking into account a multi-user system where the duration of submitted videos is an unknown variable, their processing cannot be attributed to simple queues. Imagining a scenario where a user submits a video with a duration of one hour, followed by another submission by another user of a video with a duration of 1 minute, if the processing of these videos was scheduled using a queue, the user that submitted the 1 minute video would have to wait for the processing of the 1 hour video plus the processing of his own video to be able to see it on-line. This does not translate into a good Quality of Service (QoS) to the user. The proposed solution to the problem is the implementation and testing of several scheduling techniques for *jobs* that define priorities in processing according to QoS parameters. After testing, one technique is chosen as the best solution to this problem. The main goal is to minimize the difference between the time it takes to process a video in the system and the time it would take to process the same video if there were no other videos in the queue. In order to balance the processing of videos efficiently, they are first divided into different chunks with the same duration. This allows the different chunks to be processed concurrently, being that in a case where there is no overload (there are free processing cores), videos are processed even faster than expected. The data on which the tests are performed is a collection of video files with different formats and durations. These videos were selected to best simulate the conditions that MEO's systems deal with every day. The results is a better balanced system in account to the distribution of processing power per user and a faster system due to the parallelization of the process.

Referências

[AB12] Hamid Arabnejad e Jorge Barbosa. Fairness Resource Sharing for Dynamic Workflow Scheduling on Heterogeneous Systems. *2012 IEEE 10th International Symposium on Parallel and Distributed Processing with Applications*, pages 633–639, 2012. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6280354, doi:10.1109/ISPA.2012.94.

[AB14] Hamid Arabnejad e Jorge G. Barbosa. A Budget Constrained Scheduling Algorithm for Workflow Applications. *Journal of Grid Computing*, 12(4):665–679, 2014. doi:10.1007/s10723-014-9294-7.

[AB15] Hamid Arabnejad e Jorge G. Barbosa. Multi-workflow QoS-constrained scheduling for utility computing. oct 2015. URL: http://dx.doi.org/10.1109/CSE.2015.29, doi:10.1109/cse.2015.29.

[ABAB13] Hamid Arabnejad, Jorge Barbosa, Hamid Arabnejad e Jorge Barbosa. Dynamic Scheduling of Workflows in Heterogeneous Systems. *Fair Resource Sharing for Dynamic Scheduling of Workflows on Heterogeneous Systems, 1st Edition*, 1:1–22, 2013.

[ABS14] Hamid Arabnejad, Jorge G. Barbosa e Fr??d??ric Suter. Fair Resource Sharing for Dynamic Scheduling of Workflows on Heterogeneous Systems. In *High-Performance Computing on Complex Environments*, pages 145–167. 2014. doi:10.1002/9781118711897.ch9.

[BM10] Luiz Fernando Bittencourt e Edmundo R M Madeira. Towards the scheduling of multiple workflows on computational Grids. *Journal of Grid Computing*, 8(3):419–441, 2010. doi:10.1007/s10723-009-9144-1.

[BM11] Jorge G. Barbosa e Belmiro Moreira. Dynamic scheduling of a batch of parallel task jobs on heterogeneous clusters. In *Parallel Computing*, volume 37, pages 428–438, 2011. doi:10.1016/j.parco.2010.12.004.

[Buy99] R Buyya. High Performance Cluster Computing: Architectures and Systems, Volume 1. *Prentice Hall PTR*, 82(Journal Article):327–350, 1999.

[CDS10] Henri Casanova, Frdric Desprez e Frdric Suter. On cluster resource allocation for multiple parallel task graphs. *Journal of Parallel and Distributed Computing*, 70(12):1193–1203, 2010. doi:10.1016/j.jpdc.2010.08.017.

[HCTY+12] Ad??n Hirales-Carbajal, Andrei Tchernykh, Ramin Yahyapour, Jos?? Luis Gonz??lez-Garc??a, Thomas R??blitz e Juan Manuel Ram??rez-Alcaraz. Multiple workflow scheduling strategies with user run time estimates on a Grid. *Journal of Grid Computing*, 10(2):325–346, 2012. doi:10.1007/s10723-012-9215-6.

REFERÊNCIAS

[HHW11] Chih-Chiang Hsu, Kuo-Chan Huang e Feng-Jian Wang. Online scheduling of workflow applications in grid environments. *Future Generation Computer Systems*, 27(6):860–870, 2011. URL: http://dx.doi.org/10.1016/j.future.2010.10.015, doi:10.1016/j.future.2010.10.015.

[KA99] Yu-Kwong Kwok e Ishfaq Ahmad. Static scheduling algorithms for allocating directed task graphs to multiprocessors. *ACM Computing Surveys*, 31(4):406–471, 1999. doi:10.1145/344588.344618.

[KSW10] David Karger, Cliff Stein e Joel Wein. Scheduling Algorithms. In *Algorithms and Theory of Computation Handbook (2nd edition)*, volume 2, pages 20:1–20:34. 2010. URL: http://dl.acm.org/citation.cfm?id=1882723.1882743.

[LCJY09] Ke Liu, Jinjun Chen, Hai Jin e Yun Yang. A Min-Min average algorithm for scheduling Transaction-intensive grid workflows. *Conferences in Research and Practice in Information Technology Series*, 99:41–48, 2009.

[MAS+99] M. Maheswaran, S. Ali, H.J. Siegal, D. Hensgen e R.F. Freund. Dynamic matching and scheduling of a class of independent tasks onto heterogeneous computing systems. *Proceedings. Eighth Heterogeneous Computing Workshop (HCW'99)*, 131:107–131, 1999. doi:10.1109/HCW.1999.765094.

[NS09] Tchimou N'Takpé e Frédéric Suter. Concurrent scheduling of parallel task graphs on multi-clusters using constrained resource allocations. In *IPDPS 2009 - Proceedings of the 2009 IEEE International Parallel and Distributed Processing Symposium*, 2009. doi:10.1109/IPDPS.2009.5161161.

[RA97] Robert Rönngren e Rassul Ayani. A comparative study of parallel and sequential priority queue algorithms. *ACM Transactions on Modeling and Computer Simulation*, 7(2):157–209, 1997. doi:10.1145/249204.249205.

[SB14] Altino M. Sampaio e Jorge G. Barbosa. Towards high-available and energy-efficient virtual computing environments in the cloud. *Future Generation Computer Systems*, 40:30–43, 2014. URL: http://dx.doi.org/10.1016/j.future.2014.06.008, doi:10.1016/j.future.2014.06.008.

[SZ04] R Sakellariou e Henan Zhao. A hybrid heuristic for DAG scheduling on heterogeneous systems. In *Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International*, pages 111–123, 2004. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1303065, doi:10.1109/IPDPS.2004.1303065.

[THW02] Haluk Topcuoglu, Salim Hariri e M Wu. Performance-effective and low-complexity task scheduling for heterogeneous computing. *Parallel and Distributed Systems, ...*, 13(3):260–274, 2002. URL: http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=993206, doi:10.1109/71.993206.

[XCWB09] Meng Xu, Lizhen Cui, Haiyang Wang e Yanbing Bi. A multiple QoS constrained scheduling strategy of multiple workflows for cloud computing. In *Proceedings - 2009 IEEE International Symposium on Parallel and Distributed Processing with Applications, ISPA 2009*, pages 629–634, 2009. doi:10.1109/ISPA.2009.95.

REFERÊNCIAS

[YS08] Zhifeng Yu e Weisong Shi. A planner-guided scheduling strategy for multiple workflow applications. In *Proceedings of the International Conference on Parallel Processing Workshops*, pages 1–8, 2008. doi:10.1109/ICPP-W.2008.10.

[ZS06] Henan Zhao e Rizos Sakellariou. Scheduling multiple DAGs onto heterogeneous systems. In *20th International Parallel and Distributed Processing Symposium, IPDPS 2006*, volume 2006, 2006. doi:10.1109/IPDPS.2006.1639387.