

P.O.Box 8718, Beijing 100080, China	Journal of Software, Mar. 2004,15(3):379-390
E-mail: jos@iscas.ac.cn	ISSN 1000-9825, CODEN RUXUEW, CN 11-2560/TP
<a href="http://www.jos.org.cn">http://www.jos.org.cn</a>	Copyright © 2004 by The Editorial Department of Journal of Software

# 可变负载动态反馈弹性调度模型及其算法研究

陈宇, 戴琼海

[Full-Text PDF](#) [Submission](#) [Back](#)

陈宇, 戴琼海 (清华大学 深圳研究生院 宽带网与多媒体研究中心, 广东 深圳 518055)

作者简介: 陈宇(1974-), 男, 四川大竹人, 博士, 主要研究领域为实时调度算法, 多媒体计算; 戴琼海(1964-), 男, 副教授, 主要研究领域为宽带网络, 图像处理.

联系人: 陈宇 Phn: +86-755-26036765 ext 815, E-mail: chenys@sz.tsinghua.edu.cn

Received 2003-03-13; Accepted 2003-10-08

## Abstract

ADue to the variation of the tasks' attributes, the behavior of soft real-time systems, such as multimedia application, is becoming increasingly unpredictable. Under this circumstance, the scheduling algorithms, which depend on the tasks' static attributes, can't give a usable and efficient resource allocation to those soft real-time systems. This paper presents an elastic scheduling algorithm for flexible workload. Based on sampling the total number and lost number of the task instances, this algorithm adjusts the number of task instances executing in the next sampling period to guarantee the tasks' basic QoS (quality of service) and to improve the system resource utilization and concurrency in the next sampling period. This paper analyzes the model and evaluates its performance. Simulation results show that, besides improving resource utilization, this algorithm has good stability and convergence.

Chen Y, Dai QH. Research on dynamic feedback and elastic scheduling model and algorithm for flexible workload. *Journal of Software*, 2004,15(3):379~390.

<http://www.jos.org.cn/1000-9825/15/379.htm>

## 摘要

由于工作负载的动态变化,以多媒体应用为代表的软实时系统的运行具有很大的不确定性.在这种情况下,依靠任务的静态属性进行调度分析和决策不足以为系统提供高效、实用的资源分配支持.提出一种弹性资源调度算法,该算法周期地采集系统的作业总数和作业丢失数,并以此为根据改变部分软实时任务的作业周期,以调整系统在下一个采样周期内的作业总数,达到满足任务的QoS(quality of service)、接纳尽可能多的服务请求、提高系统的并发服务能力的目的.详细分析了模型结构和核心算法的实现机制,并利用模拟平台对该算法进行了验证.实验结果表明,该算法在提高资源利用效率的同时,还具有良好的稳定性和收敛性.

## References:

- [1] Liu JWS, Shin WK, Lin KJ, Bettati R, Chung JY. Imprecise computations. *Proc. of the IEEE*, 1994,82(1):83~94.
- [2] Dey JK, Kurose J, Towsley D. On-Line processor scheduling for a class of IRIS real-time tasks. *IEEE Trans. on Computers*, 1996,45(7):217~228.
- [3] Hamdaoui M, Ramanathan P. A dynamic priority assignment technique for streams with (m,k)-firm deadlines. *IEEE Trans. on Computers*, 1995,44(12):1443~1451.
- [4] Aydin H, Melhem R, Mossé D, Mejía-Alvarez P. Optimal reward-based scheduling for periodic real-time tasks. *IEEE Trans. on Computers*, 2001,50(2):111~130.
- [5] Chen Y, Xiong GZ. Imprecise computation fault-tolerant rate-monotonic scheduling. In: Zhou WL, ed. *Proc. of the 5th Int'l Conf. on Algorithms and Architectures for Parallel Processing*. Beijing: IEEE Computer Society, 2002. 278~285.

- [6] Chen Y, Yu X, Xiong GZ. Fault-Tolerant earliest deadline first scheduling with resource reclaim. In: Zhou WL, ed. Proc. of the 5th Int'l Conf. on Algorithms and Architectures for Parallel Processing. Beijing: IEEE Computer Society, 2002. 290~293.
- [7] Kuo T-W, Mok AK. Load adjustment in adaptive real-time system. In: Jeffay K, ed. Proc. of the 12th IEEE Real-Time System Symp. San Antonio: IEEE Computer Society, 1991. 160~170.
- [8] Nakajima T, Tezuka H. A continuous media application supporting dynamic QoS control on real-time mach. In: Dittai Z, ed. Proc. of the ACM Multimedia'94. San Francisco: ACM Press, 1994. 289~297.
- [9] Lee C, Rajikumar R, Mercer C. Experiences with processor reservation and dynamic QoS in real-time mach. In: Aigrain P, ed. Proc. of the ACM Multimedia'96. Boston: ACM Press, 1996. 22~31.
- [10] Beccari G, Caselli S, Reggiani M, Zanichelli F. Rate modulation of soft real-time tasks in autonomous robot control systems. In: Burns A, ed. Proc. of the 11th Euromicro Conf. on Real-Time Systems. New York: IEEE Computer Society, 1999. 21~28.
- [11] Buttazzo GC, Lapari G, Caccamo M, Abeni L. Elastic scheduling for flexible workload management. IEEE Trans. on Computers, 2002,51(3):289~302.
- [12] Buttazzo G, Abeni L. Adaptive workload management through elastic scheduling. Real-Time Systems, 2002,23(3):7~24.
- [13] Lu CY, Stankovic JA, Son SH, Tao G. Feedback control real-time scheduling: Framework, modeling and algorithms. Real-Time Systems, 2002,23(3):85~126.
- [14] Abdelzaher TF, Bhatti N. Web server QoS management by adaptive content delivery. In: Bhatti S, ed. Proc. of the Int'l Workshop on Quality of Service. London: IEEE Computer Society, 1999. 216~225.
- [15] Abdelzaher TF, Lu CY. Modeling and performance control of internet servers. In: Bitmead B, ed. Proc. of the 39th IEEE Conf. on Decision and Control. Sydney: IEEE Computer Society, 2000. 2234~2239.
- [16] Lu CY, Abdelzaher TF, Stankovic JA, Son SH. A feedback control approach for guaranteeing relative delays in Web servers. In: Ho JM, ed. Proc. of the IEEE Real-Time Technology and Application Symp. Taipei: IEEE Computer Society, 2001. 51~62.
- [17] Parekh S, Gandhi N, Hellerstein J, Tilbury D, Jayram TS, Bigus J. Using control theory to achieve services level objectives in performance management. In: Zimmer W, ed. IFIP/IEEE Int'l Symp. on Integrated Network Management. Seattle: IEEE Computer Society, 2001. 841~854.
- [18] Hollot CV, Misra V, Towsley D, Gong WB. A control theoretic analysis of RED. In: Izmailov R, ed. Proc. of the IEEE Infocom. Anchorage: IEEE Computer Society, 2001. 1510~1519.
- [19] Cervin A, Eker J, Bernhardsson B, ?rzén KE. Feedback-Feedforward scheduling of control tasks. Real-Time Systems, 2002,23(3):25~53.
- [20] Chen Y. Research on supporting techniques for high dependable fault tolerant real-time systems [Ph.D. Thesis]. Chengdu: University of Electronic Science and Technology of China, 2002 (in Chinese with English abstract).

附中文参考文献:

- [20] 陈宇.高可靠容错实时系统的支撑技术研究[博士学位论文].成都:电子科技大学,2002.