

P.O.Box 8718, Beijing 100080, China	Journal of Software, August 2007,18(8):2027-2037
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 © 2007 by <i>Journal of Software</i>

# 虚拟计算环境中的多机群协同调度算法

张伟哲, 田志宏, 张宏莉, 何 慧, 刘文懋

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

张伟哲<sup>1</sup>, 田志宏<sup>1,2</sup>, 张宏莉<sup>1</sup>, 何 慧<sup>1</sup>, 刘文懋<sup>1</sup>

<sup>1</sup>(哈尔滨工业大学 计算机科学与技术学院, 黑龙江 哈尔滨 150001)

<sup>2</sup>(中国科学院 计算技术研究所, 北京 100080)

作者简介: 张伟哲(1976—),男,黑龙江哈尔滨人,博士,讲师,CCF会员,主要研究领域为网络计算,网络安全.田志宏(1978—),男,博士,讲师,CCF会员,主要研究领域为网络安全,网络计算.张宏莉(1973—),女,博士,教授,博士生导师,CCF会员,主要研究领域为并行计算,网络安全,拓扑发现.何慧(1974—),女,博士,副教授,CCF会员,主要研究领域为网络计算.刘文懋(1983—),男,硕士生,主要研究领域为网络计算.

联系人: 张伟哲 Phn: +86-451-86419652, Fax: +86-451-86413309, E-mail: wzzhang@hit.edu.cn, <http://www.hit.edu.cn>

Received 2007-02-24; Accepted 2007-04-26

## Abstract

Based on the core mechanisms of Internet-based virtual computing environment (iVCE), a novel architectural framework for the multi-cluster task co-allocation is proposed by introducing the autonomic scheduling elements, domain scheduling commonwealth and meta-scheduling executor. A new multi-cluster task scheduling schema based on the multi-cluster task execution performance model is presented. Four multi-cluster heuristic scheduling algorithms are provided. Experiments indicate the scheduler schema and the algorithms are effective in the objective function of makespan and average utilization.

Zhang WZ, Tian ZH, Zhang HL, He H, Liu WM. Multi-Cluster co-allocation scheduling algorithms in virtual computing environment. *Journal of Software*, 2007,18(8):2027-2037.

DOI: 10.1360/jos182027

<http://www.jos.org.cn/1000-9825/18/2027.htm>

## 摘要

基于虚拟计算环境的核心机理,提出由自主调度单元、域调度共同体、元调度执行体为核心的多机群协同系统框架.剖析多机群任务并发运行性能模型,设计了多机群协同调度算法框架,提出最大空闲节点优先、最小网络拥塞优先、最小异构因子优先与最小异构空闲节点优先4种启发式资源选择策略.实验验证了协同调度模型与算法在任务集完成时间与系统平均利用率的测度上的有效性.

基金项目: Supported by the National Natural Science Foundation of China under Grant No.90412001 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2006AA02Z334 (国家高技术研究发展计划(863)); the National Basic Research Program of China under Grant No.G2005CB321806 (国家重点基础研究发展计划(973))

## References:

- [1] Feitelson DG, Rudolph L, Schwiegelshohn U. Parallel job scheduling—A status report. In: Feitelson DG, Rudolph L, Schwiegelshohn U, eds. *Proc. of the Job Scheduling Strategies for Parallel Processing*. LNCS 3277, Berlin: Springer-Verlag, 2004. 1-16.
- [2] Dong FP, Akl SG. Scheduling algorithms for grid computing: State of the art and open problems. Technical Report, 2006. <http://www.cs.queensu.ca/TechReports/Reports/2006-504.pdf>
- [3] Abawajy JH, Dandamudi SP. Parallel job scheduling on multicluster computing systems. In: *Proc. of the IEEE Int'l Conf. on Cluster Computing (CLUSTER 2003)*. Oakland: IEEE Computer Press, 2003. 11-18.

- [4] Sabin G, Kettimuthu R, Rajan A, Sadayappan P. Scheduling of parallel jobs in a heterogeneous multisite environment. In: Feitelson DG, Rudolph L, Schwiegelshohn U, eds. Proc. of the Job Scheduling Strategies for Parallel Processing. Berlin: Springer-Verlag, 2003. 87-104.
- [5] Ernemann C, Hamscher V, Yahyapour R. Benefits of global grid computing for job scheduling. In: Buyya R, ed. Proc. of the 5th IEEE/ACM Int'l Workshop on Grid Computing in Conjunction with SuperComputing 2004. Oakland: IEEE Computer Press, 2004. 374-379.
- [6] Hamscher V, Schwiegelshohn U, Streit A, Yahyapour R. Evaluation of job-scheduling strategies for grid computing. In: Proc. of the Grid Computing (Grid 2000) at 7th Int'l Conf. on High Performance Computing (HiPC 2000). LNCS 1971, Berlin: Springer-Verlag, 2004. 191-202.
- [7] Ernemann C, Hamscher V, Streit A, Yahyapour R. Enhanced algorithms for multisite scheduling. In: Proc. of the 3rd IEEE/ACM Int'l Workshop on Grid Computing (Grid 2002) at Supercomputing 2002. LNCS 2536, Berlin: Springer-Verlag, 2002. 219-231.
- [8] Ernemann C, Hamscher V, Yahyapour R, Streit A. On effects of machine configurations on parallel job scheduling in computational grids. In: Proc. of the Int'l Conf. on Architecture of Computing Systems. Springer-Verlag, 2002. 169-179.
- [9] Sinaga JMP, Mohamed HH, Epema DHJ. A dynamic co-allocation service in multicluster systems. In: Feitelson DG, Rudolph L, Schwiegelshohn U, eds. Proc. of the Job Scheduling Strategies for Parallel Processing. LNCS 3277, Berlin: Springer-Verlag, 2005. 194-209.
- [10] Bucur AID, Epema DHJ. The maximal utilization of processor co-allocation in multicluster systems. In: Werner B. ed. Proc. of the 17th Int'l Parallel and Distributed Processing Symp. (IPDPS 2003). Oakland: IEEE Computer Press, 2003. 60-69.
- [11] Bucur AID, Epema DHJ. The performance of processor co-allocation in multicluster systems. In: Lee S, Sekiguchi S, eds. Proc. of the 3rd IEEE/ACM Int'l Symp. on Cluster Computing and the Grid (CCGrid 2003). Oakland: IEEE Computer Press, 2003. 302-309.
- [12] Bucur AID, Epema DHJ. The influence of the structure and sizes of jobs on the performance of co-allocation. In: Feitelson DG, Rudolph L, eds. Proc. of the 6th Workshop on Job Scheduling Strategies for Parallel Processing. LNCS 1911, Berlin: Springer-Verlag, 2000. 154-173.
- [13] Bucur AID, Epema DHJ. The influence of communication on the performance of co-allocation. In: Feitelson D, Rudolph L, eds. Proc. of the 7th Workshop on Job Scheduling Strategies for Parallel Processing. LNCS 2221, Berlin: Springer-Verlag, 2001. 66-86.
- [14] William J, Walter L, Louis P, Daniel S. Characterization of bandwidth-aware meta-schedulers for co-allocating jobs across multiple clusters. Journal of Supercomputing, 2005,34(2):135-163.
- [15] Ding J, Chen GL, Gu J. A unified resource mapping strategy in computational grid environments. Journal of Software, 2002,13(7): 1303-1308 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/13/1303.pdf>
- [16] Lin WW, Qi DL, Li YJ, Wang ZY, Zhang ZL. Independent tasks scheduling on tree-based grid computing platforms. Journal of Software, 2006,17(11):2352-2361 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/17/2352.htm>
- [17] Zhang WZ, Albert C, Hu MZ. Multisite co-allocation algorithms for computational grid. In: Proc. of the 3rd High-Performance Grid Computing Workshop (HPGC 2006), Associated with the Int'l Parallel and Distributed Processing Symp. 2006 (IPDPS 2006). New York: IEEE Press, 2006. 8.
- [18] Zhang WZ, Fang BX, Hu MZ, Liu XR, Zhang HL, Gao L. Multisite co-allocation scheduling algorithms for parallel jobs in computing grid environments. Science in China (Series E), 2006,36(10):1240--262 (in Chinese with English abstract).
- [19] Foster I, Kesselman C, Tuecke S. The anatomy of the grid: Enabling scalable virtual organizations. Int'l Journal of High Performance Computing Applications, 2001,15(3):200-222.
- [20] Lu XC, Wang HM, Wang J. Internet-Based virtual computing environment (iVCE): Concepts and architecture. Science in China (Series E), 2006,36(10):1081-1099 (in Chinese with English abstract).
- [21] Foster I. Service-Oriented science. Science, 2005,308(5723):814-817.
- [22] Huai JP, HU CM, Li JX, Sun HL, Wo TY. CROWN: A service grid middleware with trust management mechanism. Science in China (Series E), 2006,36(10):1127-1155 (in Chinese with English abstract).

[23] Mei H, Huang K, Zhao HY, Jiao WP. A software architecture centric engineering approach for Internetwork. Science in China (Series E), 2006,36(10):1100-1126 (in Chinese with English abstract).

[24] Buyya R, Abramson D, Giddy J, Stockinger H. Economic models for resource management and scheduling in grid computing. Special Issue on Grid Computing Environments, Journal of Concurrency and Computation: Practice and Experience (CCPE), 2002,14(13-15):1507-1542.

[25] Dail H, Berman F, Casanova H. A decoupled scheduling approach for grid application development environments. Journal of Parallel and Distributed Computing, 2003,63(5):505-524.

[26] Du XL, Jiang CJ, Xu GR, Ding ZJ. A grid DAG scheduling algorithm based on fuzzy clustering. Journal of Software, 2006,17(11): 2277-2288 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/17/2277.htm>

[27] Sodhi S, Subhlok J. Skeleton based performance prediction on shared networks. In: Moreira E, ed. Proc. of the 4th IEEE Symp. on Cluster Computing and the Grid (CCGrid 2004). Washington: IEEE Computer Press, 2004. 723-730.

附中文参考文献:

[15] 丁箐,陈国良,顾钧. 计算网格环境下一个统一的资源映射策略. 软件学报, 2002,13(7):1303-1308. <http://www.jos.org.cn/1000-9825/13/1303.pdf>

[16] 林伟伟,齐德昱,李拥军,王振宇,张志立. 树型网格计算环境下的独立任务调度. 软件学报, 2006,17(11):2352-2361. <http://www.jos.org.cn/1000-9825/17/2352.htm>

[18] 张伟哲,方滨兴,胡铭曾,刘欣然,张宏莉,高雷. 计算网格环境下基于多址协同的作业级任务调度算法. 中国科学(E辑), 2006,36(10): 1240-1262.

[20] 卢锡城,王怀民,王戟. 虚拟计算环境iVCE:概念与体系结构. 中国科学(E辑), 2006,36(10):1081-1099.

[22] 怀进鹏,胡春明,李建欣,孙海龙,沃天宇. CROWN:面向服务的网格中间件系统与信任管理. 中国科学(E辑), 2006,36(10): 1127-1155.

[23] 梅宏,黄罡,赵海燕,焦文品. 一种以软件体系结构为中心的网构软件开发方法. 中国科学(E辑), 2006,36(10):1100-1126.

[26] 杜晓丽,蒋昌俊,徐国荣,丁志军. 一种基于模糊聚类的网格DAG任务图调度算法. 软件学报, 2006,17(11):2277-2288. <http://www.jos.org.cn/1000-9825/17/2277.htm>