

网格环境下基于门限的应用监控策略^{*}

向穗华¹, 韩建军²

(1. 泉州师范学院 理工学院, 福建 泉州 362000; 2. 华中科技大学 计算机科学与技术学院, 武汉 430074)

摘 要: 为了尽可能地保证更多的网格应用在截止时间内完成, 采用模糊统计法和最大隶属度原则计算出网格应用性能差率的门限, 利用该门限对应用作出相应调整, 部分解决了现有网格监控的不足。最后进行了仿真测试, 从总运行时间和截止期内的完成率两个方面说明了该监控是有效的。

关键词: 网格监控; 模糊统计; 隶属度; 门限

中图分类号: TP391

文献标志码: A

文章编号: 1001-3695(2008)02-0575-03

Threshold-based application monitoring on grid

XIANG Sui-hua¹, HAN Jian-jun²

(1. College of Technology, Quanzhou Normal university, Quanzhou Fujian 362000, China; 2. College of Computer Science & Technology, Huazhong University of Science & Technology, Wuhan 430074, China)

Abstract: To guarantee much more applications could be accomplished before deadline, fuzzy statistics and the biggest membership degree principle were introduced to compute the threshold of difference rate of grid application performance, on which to reschedule the tasks based. Partially supplied the gap of the grid monitoring. And experimental results of total runtime and completion rate of all tasks show that the scheme is effective.

Key words: grid monitoring; fuzzy statistics; membership degree; threshold

网格资源是分布的、异构的、非独占的、数量庞大而且随机变化的。满足应用性能指标的网格资源在一段时间后, 可能变得负载繁重; 不满足条件的网格资源在一段时间后可能相当空闲。在网格节点中, 通常都涉及到具有不同优先级和截止时间的多个应用对节点资源的竞争, 所以为了使网格尽可能地保证更多的应用在截止时间内完成, 网格监控除了采集资源性能信息外, 对应用的运行性能进行监控也是很重要的。

目前虽然已经出现了大量的网格监控系统, 如 R-GMA^[2]、MDS2^[3]、MDS4^[4]、GridRM^[5]、NWS^[6]、Ganglia^[7] 和 Hawkeye^[8] 等, 但是监控网格应用程序的系统较少。GRM^[9] (grid resource monitor) 收集在异构分布式系统中运行的应用程序信息 (包括事件跟踪数据和应用程序行为的统计信息), 一旦应用程序发生错误, 用户就可以根据跟踪事件找到失效点。GRM^[10] 作了改进, 增加了两个网格服务, 即 jobmanager 和 submitter, 提供了在多层的网格监控体系中作业与进程之间如何简单而且安全地映射, 跟踪节点中的进程是否运行等。NetLogger^[11] 是分布式应用程序、主机和网络记录器, 主要应用于性能和瓶颈分析等。文献[12]是一个分布式性能测量和资源控制系统, 分布式性能传感器负责从执行的应用程序中抽取定性定量的性能数据, 为决策系统提供所需要的数据。以上系统缺乏对应用运行性能的分析, 没有提供运行性能优与劣的判断策略及相应的处理方法。本文提出了一种对网格任务运行性能优劣进行基于门限的监控, 并对应用进行了相应的调整以提高截止期内应用的完成率。

1 网格监控

网格监控主要分为两个方面, 即网格资源性能信息监控和

网格应用程序性能信息监控。网格资源信息是应用调度的基础, 它包括 CPU 利用率、等待队列长度、内存利用率和网络带宽等动态信息及 IP 地址、操作系统类型等静态信息。应用程序 (包括计算密集型、通信密集型和计算通信相对均衡型应用) 性能信息的监控和预测是应用顺利完成的保证, 它包括应用程序对 CPU 的占用率、内存的占用率、网络带宽的占用率和已执行时间等动态信息和所属用户等静态信息。

Tierney 等人^[13] 提出了网格监控体系结构 GMA。全球网格论坛 GGF 将其采纳为网格监控系统的标准体系结构。GMA 由三个主要组件构成 (图 1)。

各个节点上的传感器或监测工具收集本节点资源和应用运行的性能数据, GMA 将监控到的性能数据以带时间戳 (time-stamped) 的事件方式传递 (一个事件是一个特定结构的数据集合)。生产者收集性能数据, 是性能数据的提供者; 消费者接收性能数据, 是性能数据的接收者。生产者和消费者都向目录服务的条目注册他们的相关信息。消费者使用目录服务发现满足条件的生产者后, 直接与生产者通信而不再牵涉目录服务。这种生产者/消费者模型, 通过将用做事件发布信息的元数据抽取出来放到一个可统一访问的位置, 即目录服务, 把监控信息的收集、发布、查询和发现相分离, 具有较好的可扩展性。但是它对一些细节并没有规定, 如监控数据的类型和格式、生产者与数据源之间的接口和协议等。这留待在具体的 GMA 实例中实现, 如 R-GMA、MDS2、MDS4、GridRM 等。

本文提出的基于门限的应用监控策略, 在应用运行的节点中为应用设置自己的应用性能传感器, 定期获取该应用的运行性能数据; 但是并不将这些数据传送到生产者, 而是在本地使用门限判断该应用运行性能的优劣, 判断是否需要申请再调度

收稿日期: 2006-10-10; 修回日期: 2007-02-26 基金项目: 国家自然科学基金资助项目 (60503048)

作者简介: 向穗华 (1981-), 女, 湖北人, 硕士, 主要研究方向为网格计算 (icy93@126.com); 韩建军 (1972-), 男, 湖北人, 副教授, 主要研究方向为并行计算、调度算法。

结果用 $A_{i,j}$ 表示,统计过程中结果 $A_{i,j}$ 出现的次数用 $n_{i,j}$ 来计量。 A_i 是边界模糊的模糊子集, $A_{i,j}$ 是边界明确的普通子集。

定义 3 设应用性能差率的论域 U 的模糊子集 A_i 得到 n_i 个不同结果 $A_{i,j}$,结果 $A_{i,j}$ 的计量为 $n_{i,j}$ 次,即 $\sum_{j=1}^{n_i} n_{i,j} = n$ 。其中: $i=1,2,3;j=1,2,\cdots,n_i$ 。

模糊子集 A_i 的模糊分布是一张有区间和计量两个属性的二维表。为了便于计算硬门限和软门限,本文不将二维表中的数据用近似函数表示。

2) 模型识别 采用模糊统计法统计出各个模糊子集的模糊分布后,使用最大隶属度原则对网格应用性能差率论域中的每个元素进行模型识别,即判断该元素属于哪个模糊子集。

定义 4 设某个应用性能差率 $u \in A_{i,j}$ 的计量次数用 $u_{i,j}$ 表示,则 $u_{i,j} = \begin{cases} 0 & u \notin A_{i,j} \\ n_{i,j} & u \in A_{i,j} \end{cases}$ 。其中: $i=1,2,3;j=1,2,\cdots,n_i$ 。

定义 5 设某个应用性能差率 u 属于 A_i 的计量次数用 u_i 表示,则 $u_i = \sum_{j=1}^{n_i} u_{i,j}$ 。其中: $i=1,2,3;j=1,2,\cdots,n_i$ 。

定义 6 设函数 $\mu_{A_i}(u)$ 表示元素 u 隶属于 A_i 的程度,即 u 隶属于 A_i 的隶属度,则 $\mu_{A_i}(u) = u_i/n$ 。

元素 u 的最大隶属度原则是指将最大的 $\mu_{A_i}(u)$ 对应的模糊子集 A_i 作为 u 所属的模型。

3) 计算硬门限 α 其实质上就是找出 A_1 中所有普通子集的并集中隶属于 A_1 的最小整型元素。计算步骤如下:(a)求 A_1 中所有普通子集的并集 $[-100,\alpha^U]$;(b)求 A_1 中所有普通子集的交集 $[-100,\alpha^N]$;(c)在区间 $[\alpha^N+1,\alpha^U]$ 中找到隶属于 A_1 的最小整型元素,该元素就为硬门限 α , $[-100,\alpha]$ 也就为硬调度区间。

4) 计算软门限 β 其实质上就是找出 A_2 中所有普通子集的并集,除去属于硬调度区间 $[-100,\alpha]$ 的所有元素中隶属于 A_2 的最大整型元素。计算步骤如下:(a)求 A_2 中所有普通子集的并集,这个并集不一定是一个连续的区间;(b)求该并集与硬调度区间 $[-100,\alpha]$ 的差集;(c)在该差集中找到隶属于 A_2 的最大整型元素,该元素就为软门限 β , $[\alpha+1,\beta]$ 也就为软调度区间。

可见,如果 $(x > \beta) \cup [(x+y)/2 > \beta]$,都无须申请再调度。

3 仿真及结果分析

采用 fuzzy 统计的方法建立 A_1 、 A_2 、 A_3 的模糊分布,调查了 250 个网格用户,他们认真考虑了应用性能差率的负大、负中和负小这三个概念的含义之后,在如图 4 所示的鉴定表格中,填入整型数值。

大:	_____ % 至 _____ %
中:	_____ % 至 _____ %
小:	_____ % 至 _____ %

图 4 应用性能差率鉴定表

得到负大、负中、负小的模糊分布,使用硬门限和软门限的计算方法计算出 $\alpha=85,\beta=60$ 。

在该模拟系统中,模拟了一个由 30 个可用节点构成的网格系统,每个主机是一个计算节点,每个主机有一个处理器,配置为 Linux 2.4.18、256 MB RAM、1.7 GHz Inter CPU。对其中的一个主机 PC₀ 进行五组测试,每次测试分配不同数目的应用到该主机上,每组测试进行两轮,其中一轮不监控。不监控的

测试中,设定超过截止期的应用仍然继续运行;另一轮是对该主机调用同样数目和同样的应用,采用基于门限的监控。在基于门限监控测试中,为所有应用设置各自的应用性能传感器,每隔 60 s 采集一次相应应用的 CPU 占用率,以该应用的 CPU 占用率作为运行性能,并在每个应用的 QoS 要求中设置 CPU 占用率指标和截止期。

第一组测试用十个应用,并且每个应用的最短完成时间不超过 300 s,其他四组测试中每个应用最短完成时间都在 1 800 s 以上。这五组测试完后,从所有应用的总运行时间和截止期内完成率两个方面进行比较,测试的结果分别如图 5、6 所示。

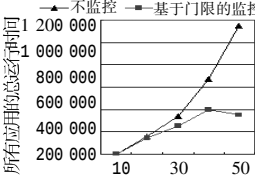


图 5 总运行时间

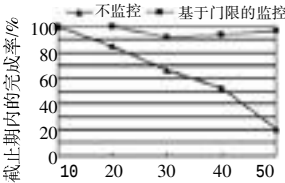


图 6 截止时间内的完成率

随着应用量的增加,基于门限监控的效果更加明显,即应用的总运行时间缩短量增加,更多的应用在截止期内完成。

4 结束语

本文介绍了对网格任务运行性能优劣的模糊监控方法。试验结果证明该方法能提高截止期内任务的完成率。在以后的工作中,进一步研究最能反映任务运行性能优劣的特征参数。

参考文献:

[1] 桂小林. 网络技术导论[M]. 北京:北京邮电大学出版社, 2005: 128-245.

[2] COOKE A, GRAY A, MA L, et al. R-GMA: an information integration system for grid monitoring[C]//Proc of the 11th International Conference on Cooperative Information Systems, Heidelberg, Herlin: Springer, 2003.

[3] CZAJKOWSKI K, FITZGERALD S, FOSTER I, et al. Grid information services for distributed resource sharing[C]//Proc of the 10th IEEE International Symposium on High Performance Distributed Computing. San Francisco: [s. n.], 2001.

[4] Globus Group. MDS4: the GT4 monitoring and discovery system[EB/OL]. [2005]. <http://www-unix.mcs.anl.gov/~schopf/Pubs/mds4.sc.pdf>.

[5] BAKER M, SMITH G. GridRM: an extensible resource management system[C]//Proc of IEEE International Conference on Cluster Computing. Hong Kong: IEEE Computer Society Press, 2003: 207-215.

[6] WOLSKI R, SPRING N, HAYES J. The network weather service: a distributed resource performance forecasting service for metacomputing[J]. Journal of Future Generation Computing Systems, 1999, 15(5-6): 757-768.

[7] MASSIE M L, CHUN B N, CULLER D E. The Ganglia distributed monitoring system: design, implementation and experience[J]. Parallel Computing, 2004, 30(7): 817-840.

[8] Hawkeye[EB/OL]. <http://www.cs.wisc.edu/condor/hawkeye>.

[9] BALATON Z, KACSUK P, PODHORSZKI N, et al. From cluster monitoring to grid monitoring based on GRM[C]//Proc of the 7th International Euro-Par Manchester on Parallel Processing. London: Springer-Verlag, 2001: 874-881.

[10] BALATON Z, GOMBAS G. Resource and job monitoring in the grid[C]//Proc of the 7th International Euro-Par Manchester on Parallel Processing. London: Springer-Verlag, 2003: 404-411.

给出了本文的仿真参数。

表 1 仿真参数

项 目	参 数
系统带宽	10 MHz
调制方式 (M_d)	BPSK/QPSK/16-QAM
符号数 (M)	30
子载波个数 (N_c)	256
导频子载波个数 (N_{ps})	4/8/16
保护间隔长度	25 个抽样间隔
扩频码长度 (SF)	16
归一化频率偏移 (Δf_c)	0.01
多普勒频移产生的归一化频率偏移 (σ_n)	$ \sigma_n \leq 0.004$
信道模型	10 路瑞利衰落信道每路有一个抽样间隔延迟

本文采用 Walsh-Hadamard 扩频码。这种码序列具有很好的互相关特性,而且在码组内所有的码序列是相互正交的^[11,12],保护间隔大于信道的最大时延扩展。在 30 个符号内有一个前导码(preamble)作有效信道估计。影响本文算法性能的主要因素是导频子载波中符号的分割块数 B。图 4~6 分别为 BPSK、QPSK、16-QAM 三种调制方式下分割块数 B 不同的估计性能误码率与信噪比的性能比较图。可以看出,误码率随着分割块数的增大性能随之改善。

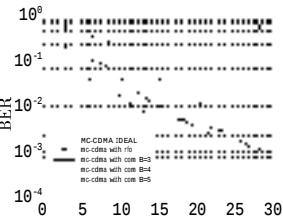


图 4 BPSK 调制 B=3、4、5 时的误码率性能比较

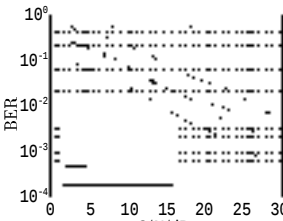


图 5 QPSK 调制 B=3、4、5 时的误码率性能比较

图 7 为 QPSK 调制模式下采用不同导频子载波数进行 RFO 估计的误码率与信噪比性能图。可以看出,导频子载波数对于误码率的性能影响不大。所以,在保证频带利用率的情况下,可采用较少的导频子载波进行估计,估计效果不会受到太大影响。

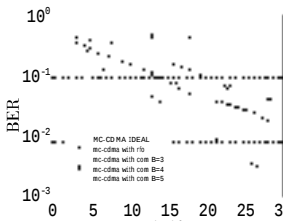


图 6 16-QAM 调制 B=3、4、5 时的误码率性能比较

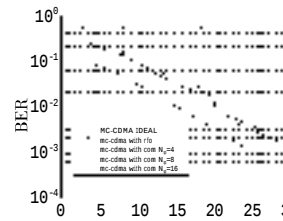


图 7 QPSK 调制 B=4 时不同导频子载波数的误码率性能比较

4 结束语

本文考虑了移动环境下多径信道传输中多普勒频移对每

个子载波的影响,在接收端频域内利用导频符号分块估计 RFO 并在时域内补偿的方法。由以上仿真结果可以看出,本文所述方法能够对接收信号的 RFO 作有效估计及补偿,对系统误码率有一定的改善。由于利用导频符号势必会降低系统的利用率,在本文研究 RFO 的估计方法基础上,还可以利用直接判决法分块估计 RFO。更多更完善的研究将能够进一步提高 RFO 估计补偿方法的准确度以及简化处理过程的复杂度。

参考文献:

[1] FAZEL K, PAPKE L. On the performance of convolutionally-coded CDMA/OFDM for Communication system[C]//Proc of IEEE Personal, Indoor and Mobile, Radio Communications. Yokohama; [s. n.], 1993:468-472.

[2] YEE N, LINNARTZ J P, FETTWEIS G. Multicarrier CDMA in indoor wireless radio networks[C]//Proc of IEEE Personal, Indoor and Mobile, Radio Communications. Yokohama; [s. n.], 1993:109-113.

[3] CHIEN F T, HWANG C H, JAYKOU C C. Performance analysis of multicarrier CDMA systems with frequency offsets and random spreading under optimum combining[J]. IEEE Trans on Communications, 2006,54(4):737-747.

[4] TOMBA L, KRZYMIEN W A. Effect of carrier phase noise and frequency offset on the performance of multicarrier CDMA systems[C]//Proc of the IEEE International Conference on Communications, 1996:1513-517.

[5] KOBAYASHI H. A novel coherent demodulation for M-QAM OFDM signal operating in the burst mode [C]//Proc of IEEE Vehicular Technology Conference. 2000:1387-1391.

[6] POLLET T, BLADEL M van, MOENECLAHEY M. BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise [J]. IEEE Trans on Communication, 1995,43(2~4):191-193.

[7] HAT K, SEA Y C, LEE C W. On a frame synchronization scheme for MC-CDMA systems in up-link asynchronous channels[C]//Proc of IEEE Vehicular Technology Conference. 1999:193-197.

[8] EGASHIRA N, TAKAYAMA H, SABA T. Improvement of CCI compensation accuracy using feedback phase tracking in MIMO-OFDM system[C]//Proc of Global IEEE Communications Conference. 2004:923-927.

[9] NEE R Van, PRASAD R. OFDM for wireless multimedia communication[M]. [S. l.]: Artech House, 2000.

[10] ZELST A Van, SCHENK T C W. Implementation of a MIMO-OFDM based wireless LAN system [J]. IEEE Trans on Signal Processing, 2004,52(2):483-494.

[11] KAISER S. OFDM-CDMA versus DS-CDMA performance evaluation for fading channels[C]//Proc of IEEE International Conference on Communications. Seattle; [s. n.], 1995:1722-1726.

[12] MILLER T, ROHLING H, GRUMHEID R. Comparison of different detection algorithms for OFDM-CDMA in broadband Rayleigh fading[C]//Proc of the 45th Vehicular Technology Conference. 1995:835-838.

trol of distributed applications[C]//Proc of the 7th IEEE International Symposium on High Performance Distributed Computing. Washington DC:IEEE Computer Society, 1998.

[13] TIERNEY B, AYDT R, GUNTER D, et al. A grid monitoring architecture, GWD-PERF16-12[R]. [S. l.]: Grid Forum Working Group, 2002.

(上接第 577 页)

[11] GUNTER D, TIERNEY B, CROWLEY B, et al. NetLogger: a toolkit for distributed system performance analysis[C]//Proc of the 8th International Symposium on Modeling, Analysing and Simulation of Computer and Telecommunication System. California; [s. n.], 2000.

[12] RIBLER R, VETTERJ S, SIMITCI H, et al. Autopilot: adaptive con-