

# 一种改进 QoS 的网格资源多维性能调度算法

焦合军<sup>1</sup>, 赵群力<sup>1</sup>, 施进发<sup>2</sup>

(1. 河南工程学院 计算机科学与工程系, 河南 郑州 451191;

2. 郑州航空工业管理学院, 河南 郑州 450015)

**摘要:** 网格环境下, 由于资源的异构性和动态性, 任务调度已扩展为多个任务在位于不同节点的异构资源上调度, 任务调度的性能直接影响到计算网格的服务质量。为提高任务调度质量, 在构建网格资源映射模型的基础上, 结合资源多维性能, 提出了一种改进 QoS 的网格资源多维性能调度算法 MQMPGR, 并且给出了与模型相对应的进行任务调度所需要的算法伪码。通过在 Gridsim 环境下的分析与比较, 仿真结果证明其优于传统的任务调度算法。

**关键词:** 网格; 服务质量; 任务调度; Gridsim 模拟器

**中图分类号:** TP393

**文献标识码:** A

**文章编号:** 1673-629X(2011)04-0101-04

## Scheduling Algorithm of Modified QoS Based on Multidimensional Performance in Grid Resources

JIAO He-jun<sup>1</sup>, ZHAO Qun-li<sup>1</sup>, SHI Jin-fa<sup>2</sup>

(1. Department of Computer Science & Engineering, Henan Institute of Engineering, Zhengzhou 451191, China;

2. Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou 450015, China)

**Abstract:** In grid environment, task scheduling has been extended to multiple tasks in different nodes for the scheduling of heterogeneous resources. The performance of task scheduling directly affects grid QoS. Combined with multidimensional performance of grid resources, to improve the quality of task scheduling, presents an algorithm for modified QoS based on multi-dimensional performance in grid resources, and gives the pseudo code description of the algorithm. This algorithm is tested on the GridSim simulator, and the results prove that it is better than traditional algorithms.

**Key words:** grid; quality of service; task scheduling; Gridsim simulator

## 0 引言

任务调度是整个网格系统的控制中枢, 在分布式和多处理机领域已有扩展性的研究。计算网格是企业有限资源条件下为了应对经济全球化、信息和服务网格化的挑战而采取的一种先进模式。随着网格研究的不断深入, 资源的广泛性和异构性进一步扩大, 用户对系统服务质量(QoS)提出了越来越高的要求, 实时任务通常采取独占方式访问共享资源, 资源的独占性会导致高优先权任务运行时往往被低优先权任务阻塞, 从而产生优先权反转, 难以满足任务的实时性。

在网格环境中, 由于资源的异构性和动态性, 任务调度已扩展为多个任务在位于不同节点的异构资源上调度。另外依据任务的重要程度不同, 网格系统用户在花费预算和运行时间上也存在多样性的需求, 比如

有的用户要求兼顾花费和时间, 任务仍要尽快完成。近年来在资源管理和调度系统问题的研究中, 网格环境中对 QoS 的要求逐渐受到关注, 文献[1]提到的组合预测机制只是在 QoS 约束中进行了横向扩展, 该工作基于 Senior 算法, 时间复杂度高。文献[2]将任务的 Deadline 约束引入 QD-Sufferage, 目的在于利用 QoS 优化网格 makespan、吞吐率等方面的综合性能, 但没有考虑用户的需求, 不能很好地解决复杂的实际网格应用问题。

文中为达到最优的资源利用率和提高任务调度性能, 在运用相关映射策略的前提下, 综合资源多维性能, 在约束参数上实现资源与任务的映射与调度, 提出了一种改进 QoS 的网格资源多维性能调度算法(algorithm for modified QoS based on multidimensional performance in grid resources, MQMPGR)。

理论分析与实验表明, 该算法提高了资源与任务的匹配效率, 减少了任务完成时间, 提高了任务调度性能。

收稿日期: 2010-08-25; 修回日期: 2010-11-17

基金项目: 国家自然科学基金项目(70971120)

作者简介: 焦合军(1981-), 男, 河南辉县人, 助教, 研究方向为网格计算, CIMS。

## 1 网格计算中与资源调度相关的因素

### 1.1 映射策略

任务映射是网格处理环境中的一个重要环节, 一个好的任务映射机制能很好地实现任务的合理分配及资源的有效利用, 从而保证网格中各部件的协同工作, 为用户提供良好的服务。映射的方法很多, 如贪婪算法<sup>[3]</sup>、超图理论、小世界理论、动态规划等。动态规划已经有了广泛的研究和应用, 是一种很成熟的算法, 满足用户 QoS 调度的要求, 文中采用动态规划模型来实现任务和资源之间的映射。

网格任务调度的实质是将相互独立的任务(假设为  $n$ ) 分配到  $m$  个异构的可用资源上, 使总的评价体系得到最优, 并使资源充分得到利用。动态规划将复杂的问题分解成多阶段过程, 在每个阶段都做出一个看上去最优的决策(在一定标准下)。决策具有无后效型, 一旦做出, 就不可再更改。为了达到最优化的目标采用的解决方案有: 对于网格中不断变化的任务需求都能够及时得到资源信息; 采取以用户为中心的策略调度以及对大量任务有效处理。所以优化网格用户的总代价关键在于整体考虑优先级和参数布局, 整合两个因素, 促使综合评价价值最优。

### 1.2 评价指标的建立

在任务调度过程中, 传统批调度的评判标准是 makespan(即第一个任务开始执行时间到最后任务结束时间之间的差值), 并未从用户的角度设置评价标准。服务质量可以满足这种需求: 对资源进行选择调度时, 引入 TQCS 准则, 依据所提出的多重 QoS 约束参数, 确定如下 6 个因子: 处理时间、资源耗费、Machine 数量、PE (Processor Element) 数量、PE 处理能力和网络带宽。所定的 6 个因子中, 时间因子就对应 TQCS 准则中的 Time, 资源耗费对应 TQCS 准则中的 Cost, PE 数量与 PE 处理能力对应 TQCS 准则中的 Service, 而网络带宽因子和 Machine 数量则对应 TQCS 准则中的 Quality。对不同的应用程序, 用户对各个因子的要求不一样, 任务之间存在不同的优先级。

各个因子对应的值大小差异很大, 对以上 6 个因子进行分类, PE 数量、PE 处理能力和网络带宽属于第一类(正向定量指标), 即越大越优指标, 以所有候选资源同类指标的最大值为理想值; 其他的属性属于第二类(负向定量指标), 即越小越优指标, 以所有候选资源同类指标的最小值为理想值。不同的指标从不

同的角度反映了资源的能力, 指标之间无法进行比较, 因此为了统一评价的有效性, 需要对各指标值进行无量纲化处理。

### 1.3 映射模型的构造

根据模型的设计原则, 得到如图 1 所示的系统框架。

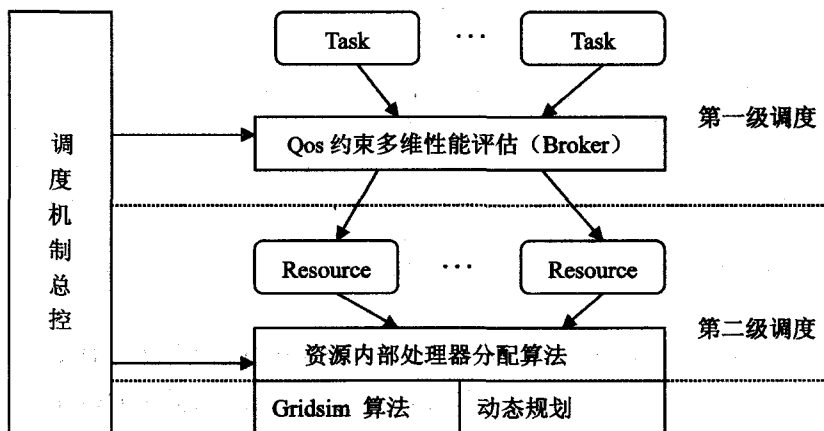


图 1 网格计算中任务调度模型

第一次调度由资源代理 QoS 约束负责, 主要优化用户对资源的要求和目标, 并将所需的资源按照一定的算法分配给用户。

第二次分配由资源实体负责, 资源实体根据内部分配算法为到达的作业分配 PE<sup>[4]</sup>。

文中采用动态规划算法作为调度算法, 与 Gridsim 环境下的算法进行分析和比较。

## 2 模型对应的算法伪码

调度算法的前提在于用户提交的任务相互独立<sup>[5]</sup>, 根据任务特性和各资源的特性, 在分配资源的第一次决策利用式(2)计算各资源的评价值, 按顺序从资源列表中选择评价值最优的资源, 并将该任务分配给该资源上。

$$\alpha + \beta + \varepsilon + \eta = 1, \alpha, \beta, \varepsilon, \eta \in [0, 1] \quad (1)$$

其中,  $\alpha, \beta, \varepsilon, \eta$  分别代表用户对任务执行时间、可靠性因子、花费和信誉度的影响因子值。用户可以根据自己对上述 4 个因素的重要性需求进行设置。利用综合评价目标函数产生决策, 根据用户设定的 TQCS 影响值, 综合考虑 6 个因子, 对资源进行评价。其值越大, 说明该资源的评价值越高、分派策略越好。

$$\text{Max} z = \alpha f_1 + \beta f_2 + \varepsilon f_3 + \eta f_4$$

$$f_k = \text{opt}_{x_k} v_k + f_{k+1}$$

$$f_{k+1} = 0, k = n, \dots, 1$$

其中, 阶段  $k = 4, 3, 2, 1$  依次表示 TQCS 四个阶段选择的过程; 决策  $x_k$  表示  $k$  阶段初可能选择的路; 阶段指标  $v_k$  表示  $k$  阶段与所选择的路段相应的量化指标。

根据现实的情况,将要研究的调度算法的任务描述如表1所示。

表1 任务请求的数据结构

数据结构参数	所需数据项
Details	id, length, 传输数据量, 输出数据量, deadline
资源 Requirements	任务需求的资源能力数, 影响因子值
State	闲置, 待分配, 待调度, 等待, 执行, 完成
Constraints	初始值, 服务方法, 调度算法(默认)

对于任意一个处在任务池中符合以上任务描述的任务,任务调度引擎将以可用资源为基础, MQMPGR 算法的任务调度过程如下:

MQMPGR ()

Initial ();

//对相关参数进行初始化,含有对任务按长度降序排列 for(每一个任务 Taski)

{  
if(满足指定的资源需求)

按照资源执行时间递增排序得到可用资源列表;

if(无满足要求的资源节点)

delete Taski with no matching Resource;

if(满足要求的资源能力数>1)

在这些同类资源中选择负载较轻的资源节点 m;

else

place node j in ListResSche;

{

else

{

按照约束条件中默认算法将任务分配到 ListResSche

allocateTaski to node j accept to the default policy;

}

}

}

上述算法中,首先对任务调度的相关参数初始化,包括降序排列任务长度,建立调度模型,之后在排列结果的基础上,将任务与资源匹配并执行调度。由算法看出,具有预定服务资源要求的任务直接与符合该需求的可用资源相匹配,并从可用资源中找到性能最好且资源动态评价最高的资源并映射到该资源,这个资源节点记录在待分配队列 ListResSche 中<sup>[6]</sup>;而没有具体资源需求要求的任务首先找到约束条件中的默认因子值,然后从候选队列中找到性能最好且负载较轻的资源并匹配该资源。对于满足条件的多个可用资源,优先选择负载低的进行映射,从而有利于网格资源均衡负载<sup>[7]</sup>。由于以调度优先级为基础,且可用资源按性能排序,因而能够有效地缩短查找资源的过程,从而缩短了任务调度。

### 3 实验仿真与结果分析

为了研究网格环境中的任务全局调度,文中用 Gridsim 模拟器模拟网格环境。采用两级调度,用户的所有任务都通过启发调度发往合适的资源,并采用 Gridsim 自带的 TimeShared 和 SpaceShared 实现本地调度。

仿真过程中,可用异构资源数为4,  $R_1, R_2, R_3, R_4$ , 资源参数见表2,假设中有10个任务要执行,在 Gridsim 下实现该算法<sup>[8]</sup>,调度过程如表3。

表2 资源参数

Resource ID	Machine 数量	资源开销 (cent/s)	PE 数量	PE 处理能力	网络带宽 (kbps)
GridResource_0	3	5	3	200 MIPS	1000
GridResource_1	2	8	6	300 MIPS	2000
GridResource_2	2	3	5	500 MIPS	3000
GridResource_3	1	6	2	600 MIPS	5000

表3 调度列表

任务 ID	任务长度/MI	开始时间/s	可用资源	完成时间/s	耗费/cent
0	9000	75.00	$R_4$	120.00	270.0
1	8000	19.06	$R_3$	59.06	120.0
2	7000	69.00	$R_1$	105.00	179.9
3	6000	59.06	$R_3$	89.06	90.0
4	5000	120.00	$R_4$	145.00	150.0
5	5000	89.06	$R_3$	114.06	75.0
6	4000	114.06	$R_3$	129.06	45.0
7	3000	105.00	$R_1$	115.00	50.0
8	2000	25.06	$R_2$	35.06	80.0
9	2000	105.00	$R_1$	150.00	225.0

10个任务分别被分配到资源  $R_2, R_3, R_3, R_4, R_1, R_2, R_4, R_4, R_3, R_1$  上执行,总运行时间为150.00,总耗费为1284.9cent。

我们分三次进行了模拟实验,在相同条件下,分别指定任务的调度算法为 MQMPGR, TimeShared 和 SpaceShared<sup>[9]</sup>,在模拟的过程中,上面提到的任务描述方法和算法伪码都得到了运用。结果如堆积数据点折线图2所示和数据点折线图3所示,其中,任务数逐步递增,步长为5。如图2所示,各种算法获得的时间随着任务数的增多而明显上升,其中 TimeShared 与 SpaceShared 获得的 Makespan(任务完成时间)相差不大,TimeShared 略优于 SpaceShared, MQMPGR 算法的调度结果明显优于 TimeShared 与 SpaceShared,体现出其更优越的调度性<sup>[10]</sup>。由图3可知,在相同任务请求下,对于不同因子,当任务数变大,用户任务开始时间要求越小时<sup>[11]</sup>,任务的执行时间相应变大,且用户代价相应变大;在任务数量很多的情况下,也适合使用 TimeShared 算法。

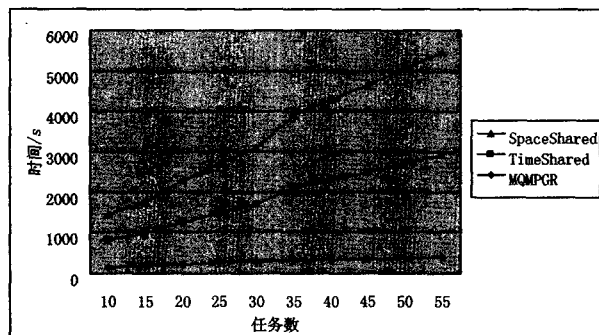


图 2 采用不同算法时对时间的影响

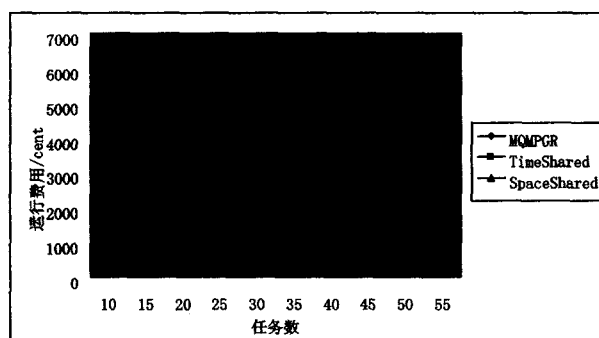


图 3 采用不同算法时对费用的影响

以往进行任务调度的时候,考虑最多的是系统吞吐率,而很少考虑用户自身的需求。总结上面描述的任务调度模型及算法伪码,可以看出,这是一个以用户为核心的调度模型,用户可以根据资源的多维性能为自己的重要性设置不同影响因子得到最优化的资源分派策略<sup>[12]</sup>,得到比较高的用户满意度。

#### 4 结束语

面向任务和资源调度,基于改进的服务质量建立任务模型和网格服务资源映射模型,根据节点资源的多维性能特性并结合动态规划理论进行资源评价,在资源评价基础上完成任务的映射与调度,提出了一种改进 QoS 的网格资源多维性能调度算法 MQMPGR。理论分析与实验结果表明,算法通过资源综合评价,有效提高了网格系统的利用率和性价比。下一步工作的重点是对调度模型在均衡分配、安全性能等方面作进一步扩展和完善,并在模型基础上研究执行时间预测机制以及提出新的更有效的调度策略,提高算法的移

植性。

#### 参考文献:

- [1] 陈 晶,孔令富,潘 勋. 结合预测机制和 QoS 约束的网格资源调度算法的研究[J]. 计算机研究与发展,2008,45(S):11-16.
- [2] 东 方,罗军舟,宋爱波. 面向 Deadline 约束的网格 Qos 任务调度[J]. 华中科技大学学报(自然科学版),2006,34(9):185-188.
- [3] 陈玉兰,郑 骏,胡文心. 一种多 QoS 约束的网格资源调度算法[J]. 华中师范大学学报(自然科学版),2009(1):111-116.
- [4] 王天擎,谢 军,曾 洲. 网格计算中的资源调度不确定性分析[J]. 计算机应用,2007,27(2):489-491.
- [5] 周 维,罗 泽,南 凯,等. 以策略为机制的网格任务调度模型研究[J]. 计算机工程,2007,33(7):89-91.
- [6] 张然美,杨寿保,申 凯,等. 网格环境下一种 QoS 感知的批调度算法[J]. 小型微型计算机系统,2007,28(6):969-973.
- [7] 陈志刚,杨 博. 网格服务资源多维性能聚类任务调度[J]. 软件学报,2009,20(10):2766-2775.
- [8] 赵 健. 基于 GridSim 的 A-MM 调度算法模拟[J]. 计算机技术与发展,2008,18(10):96-98.
- [9] Freund R F, Gherrity M, Ambrosius S, et al. Scheduling resources in multi-user, heterogeneous, computing environments with Smart-Net[C]//In: Proc The 7th IEEE Heterogeneous Computing Workshop (HCW'98). Orlando, Florida, USA: [s. n.], 1998:184-199.
- [10] Kim J K. Dynamic Mapping in fl Heterogeneous Environment with Tasks Having Priorities and Multiple Deadlines[C]//Heterogeneous Computing Workshop. Nice, France: [s. n.], 2003.
- [11] Shi Jinfa, Jiao Hejun, Sun Jianhui. Research on Collaborative Design System of small and medium-sized enterprises for Networked Manufacturing[C]//in Proc. 38th International Conference on Computers and Industrial Engineering. Beijing, China: [s. n.], 2008:2146-2153.
- [12] Zomaya A Y, Wright M. Observation on Using Genetic-Algorithms for Channel Allocation in Mobile Computing[J]. IEEE Transactions on Parallel and Distributed System, 2002, 13(9):948-962.

(上接第 100 页)

- [8] 周 曲,颜国正,王文兴,等. 基于倒谱分析的散焦模糊图像参数识别[J]. 武汉大学学报,2008,33(3):318-321.
- [9] 田 涛,潘俊民. 基于散焦图像深度测量的一种新方法[J]. 机器人,2001,23(1):15-19.
- [10] 张燕妮. 一种离焦图像复原的改进方法[D]. 大连:大连理工大学电子与信息工程学院,2005.
- [11] 张 航,罗大庸. 图像盲复原算法研究现状及其展望[J]. 中国图象图形学报,2004,9(10):1145-1152.
- [12] Krahmer F, Lin Youzuo, McAdoo B, et al. Blind Image Deconvolution: Motion Blur Estimation[M]// In IMA Preprints Series, 2006.