

基于拍卖机制的网格作业调度遗传算法

穆晓芳, 赵月爱, 张朝霞

(太原师范学院计算机系, 太原 030012)

摘要: 根据拍卖机制的在线信誉网格资源管理模型, 设计基于遗传算法的网格作业全局调度策略。引入时间和花费2个经济因素, 改进遗传算法的收敛性。改进算法根据用户需求得到最优化的资源分派策略, 提高网格系统的总体性能。将该算法纳入 GridSim 模拟器进行系统测试, 结果证明其优于传统调度算法。

关键词: 网格; 作业调度; 拍卖模型; 遗传算法; GridSim 模拟器

Genetic Algorithm for Job Scheduling in Grid Based on Auction Mechanism

MU Xiao-fang, ZHAO Yue-ai, ZHANG Zhao-xia

(Department of Computer, Taiyuan Normal University, Taiyuan 030012)

【Abstract】 According to an online reputation grid resource management model of auction mechanism, this paper designs the global scheduling strategy of grid jobs based on the Genetic Algorithm(GA). It introduces two economy factors which are time and cost, and improves GA's astringency. The improved algorithm can achieves the optimized resource allocation strategy on user demand, and improves the system's performance. This algorithm is tested on the GridSim simulator, and the results prove that it is better than traditional algorithms.

【Key words】 grid; job scheduling; auction model; Genetic Algorithm(GA); GridSim simulator

网格计算属于分布式计算, 网格资源在广域上分布, 本质上异构, 且归属不同的人和组织拥有、具有相异的存取和花费模式, 其负载和可用性动态变化, 因此, 针对传统分布式计算环境设计的资源调度策略在网格环境中效果不理想。本文使用经济方法对网格资源进行管理分配^[1], 利用经济学中的竞争理念, 形成一个平等健壮的交易市场。笔者提出的网格作业全局调度策略是基于拍卖机制的在线信誉网格资源管理模型^[2], 采用拍卖机制管理资源, 并在网格作业全局调度中引入遗传算法(Genetic Algorithm, GA)。

1 网格作业调度

网格作业调度是一种网格资源使用形式, 它根据用户确定的流程, 为用户提供使用资源的功能。网格作业一般都在远程节点上运行, 作业提交者对远程设备的控制能力很有限, 因此, 为了有效管理作业的运行, 需要网格作业管理机制, 用于管理整个网格作业的运行。

在网格计算环境中, 用户希望将各个计算站点的计算能力进行“累加”, 形成一个虚拟的超级计算站点。由于现有研究水平和网络条件不能有效支持跨站点分布运算作业, 因此在本文构筑的计算网格模型中, 不考虑计算作业的跨站点执行, 资源消费者以作业作为网格调度器进行任务分配的最小单位, 大型计算作业的分解由资源消费者完成。作业间没有相关性, 避免了作业间的频繁通信造成的性能下降。

网格作业调度的实质是将 n 个相互独立的任务分配到 m 个异构的可用资源上, 使总的任务完成时间最小, 并使资源得到充分利用。每个资源节点的速度可以表达为每单位时间的循环数。每个任务的长度可以表达为循环次数, 即任务 T_n 的执行需要 P_j 个循环, 而资源 R_m 的处理速度为 C_i 循环/s, 设 F_j 为任务 j 的最后完成时间, 跨度 makespan 为 $F_{\max} = \max\{F_j,$

$j=1,2,\dots,N\}$ 。调度任务转化为在 $2m$ 个可能的资源子集空间中寻找最优集合, 使跨度 F_{\max} 和 ΣF_j 最小。它是一个 NP 完全问题, 表明了网格的启发性本质^[3]。

2 基于GA的网格作业全局调度策略

2.1 基于GA的网格作业全局调度策略设计思路

GA 是一种普适的全局优化算法, 但标准 GA 没有很好地利用问题的领域知识, 其收敛速度通常较慢, 且其局部求精能力较弱。如何在求泛和求精间找到最佳平衡点, 是 GA 获取高性能的关键^[4]。本文 GA 基于经济模型网格, 在调度中引入经济因素, 考虑如何在满足用户条件约束的情况下取得最大收益, 可以弥补 GA 在局部求精方面的不足, 取得了更好的调度结果。本文主要考虑用户的预算花费和时间期限 2 个因素, 用户根据自己对不同经济因素的关心程度为每个因子设置不同值, 并通过两者的数学关系给出最优资源分派策略。

对网格中作业调度的 2 个阶段, 即资源分派和元调度, 本文算法考虑第 1 阶段, 即资源分派。在 GridSim 中模拟时, 资源分派使用本文算法, 元调度利用 GridSim 自带的本地调度机制。

2.2 GA的性能评估

2.2.1 性能评估标准

假设在基于拍卖机制的在线信誉资源管理模型中, 使用 GA 实现用户作业群的全局调度。其根本目标是用户的综合

基金项目: 山西省青年自然科学基金资助项目(2008021025)

作者简介: 穆晓芳(1974-), 女, 讲师、硕士, 主研方向: 分布式计算, 网络安全, 图像处理; 赵月爱, 讲师、博士研究生; 张朝霞, 讲师

收稿日期: 2008-11-10 **E-mail:** mu_xiaofang@yahoo.com.cn

评价值最优,因此,评估调度算法性能基于如下2条规则:

- (1)用户作业群的总执行时间最少;
- (2)用户作业群的总花费最低。

要使用户所有作业的总执行时间最少,应尽可能把作业分配给 PE(Processor Element)处理速度快、PE 数量多的资源节点。而此类资源的处理花费很高,因此,采用上述方式虽然可以使处理作业总执行时间最少,但其处理花费通常偏高。若把作业发往处理花费较低的资源,虽然会降低处理花费,但会增加处理总时间。因此,优化网格用户总代价的关键是综合考虑时间和花费,整合2个因素,使综合评价值最优。

2.2.2 解决方案

为了达到最优化目标,本文采用如下解决方案:

- (1)在 GA 适应值的计算过程中,整合作业的执行时间和花费2个因素。
- (2)用户可以灵活地改变执行时间和花费之间的比重。

2.3 算法描述

算法的前提是假设用户提交的作业相互独立。在分配作业前,根据作业属性(长度)和各个资源的属性(执行代价和计算能力),在 GA 算法的一代种群中利用式(2)计算各个资源的评价,依次从资源列表中选择评价值最优的资源,并将该作业分配到该资源上。

$$A+\beta=1(\alpha,\beta\in(0,1)) \quad (1)$$

其中, A 代表用户对作业花费的影响因子值; β 代表用户对作业执行时间的影响因子值。用户可以根据自己对上述2个因素的重要性需求对它们进行设置。

$IAD(Integrated Appraisal Degree)$ 代表在该资源上执行该任务时,用户的综合评价值。它根据用户设定的影响因子综合考虑2个因素,对资源进行评价,其值越小,说明该资源的评价值越高、分派策略越好。

$$IAD = \alpha \cdot (budget_resource / budget_user) + \beta \cdot (deadline_resource / deadline_user) \quad (2)$$

其中, $budget_resource$ 表示在某个资源上运行该作业产生的花费; $budget_user(budget_user \neq 0)$ 表示用户指定的预算花费; $deadline_resource$ 表示某个资源完成该作业的时间; $deadline_user(deadline_user \neq 0)$ 表示用户作业指定的时间期限。

2.4 改进的GA设计

下文分别介绍添加了时间和花费后,改进的 GA 编码表示、初始种群的产生、适应值的计算、遗传算子的设计和算法框架。

2.4.1 染色体编码表示

GA 的染色体采用间接编码方式,每个染色体是一个作业分配的优先队列 q , 其中的每个元素 q_i 是待调度的作业。每个作业必须在 q 中出现一次,作业在 q 中的次序表示它们的调度顺序。因为染色体并不是调度,所以需要解码算法把染色体映射为有效的调度。为了获得总体最优评价值,GA 采用顺序最优算法把 q 解码为有效调度,即对所有作业,按它们在 q 中的顺序,依次为其选择具有最佳评价值的资源节点。

2.4.2 初始种群

初始种群采用随机算法生成,先把所有作业放入一个集合,然后不断从集合中随机取出一个作业,把它依次放入优先队列 q 中,直到作业集为空,即所有作业都进入 q 中为止。

2.4.3 适应值函数

适应值是用于评价解的优劣,本文认为一个调度的总评

价值越小,其适应值越好。设 $MaxIAD$ 为一代种群中最差个体的评价值, $MinIAD$ 为全局最佳个体的评价值,定义 GA 中第 i 个染色体 q_i 的适应值函数为

$$Fitness(q_i) = \frac{MaxIAD - IAD(q_i) + r}{MaxIAD - MinIAD + r} \quad (3)$$

其中, r 是用来调整个体分布概率的参数。

2.4.4 遗传操作算子

标准 GA 的操作算子一般包括选择、交叉和变异3种基本形式,它们模拟了自然选择和遗传过程中发生的繁殖、杂交和突变现象,是使 GA 具备强大搜索能力的核心。

选择操作是 GA 中环境对个体适应性的评价方式,是实现群体优良基因传播的基本方式,它保证了 GA 迭代中适者生存、优胜劣汰的群体进化现象。本文 GA 中的选择算子采用轮盘赌选择方式,定义个体 q_i 被选中的概率为

$$p(q_i) = Fitness(q_i) / \sum_{j=1}^{PopSize} Fitness(q_j) \quad (4)$$

其中, $PopSize$ 为种群大小。

在 GA 中,杂交算子是最主要的搜索算子,它模仿自然界有性繁殖的基因重组过程,作用是将原有优良基因遗传给下一代个体,并生成包含更复杂基因结构的新个体。本文 GA 中的交叉算子采用一点交叉,在交叉过程中,为了保证所产生后代的合法性,先随机产生一个交叉点,然后使交叉点左边的任务顺序保持不变,交叉点及其右边的任务按它在另一个父本中的顺序进行重新排序,从而保证交叉操作后,每个个体都是合法的。染色体的交叉操作如图1所示。

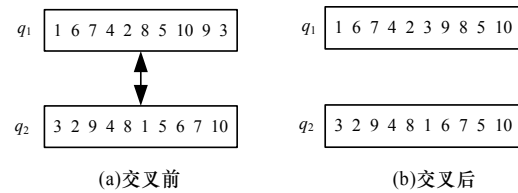


图1 染色体的交叉操作

变异算子的主要用途是拓展新的搜索空间,当种群局部收敛时,通过变异算子的突变作用来保持一定的种群多样性。本文 GA 中的变异算子随机选择2个 q 中的位置,并交换这2个位置上的任务。染色体的变异操作如图2所示。

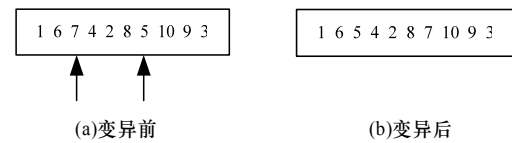


图2 染色体的变异操作

2.4.5 改进的 GA 算法框架

本文 GA 中的精英策略能保证最优解适应值的单调递增性,算法步骤如下:

- (1)采用随机算法生成初始种群。
- (2)利用顺序最优算法把染色体映射为调度并计算个体的适应值。
- (3)不满足结束条件时,重复以下操作:
 - 1)施加选择算子;
 - 2)以既定概率施加交叉算子;
 - 3)以既定概率施加变异算子;
 - 4)利用顺序最优算法把染色体映射为调度并计算个体的适应值;

5)采用精英策略保留全局最佳个体。

3 基于GA的全局作业调度算法测试

为了研究网格环境中的作业全局调度,本文用 GridSim^[5] 模拟器对网格环境进行模拟。GridSim 通过资源的“买”和“卖”引入经济模型,从而达到控制网格资源使用的目的。

3.1 算法测试说明

主要测试基于 GA 的作业全局调度策略性能,采用 2 级调度。用户的所有作业都通过 Broker 调度发往合适资源,并采用 GridSim 自带的 Timeshared 策略实现本地调度。

GridSim 本身提供了代价最优算法、时间最优算法,测试本文 GA 和 GridSim 自带的算法,并对结果进行对比分析。

3.2 算法测试结果与分析

测试基于 GA 的调度策略对网格用户综合评价价值的影响。测试时,先创建 50 个作业,分别采用基于 GA 的调度策略和 GridSim 提供的调度策略将作业分配给资源执行,GA 中的 2 个因子设定为 0.5。运行参数设置如下:种群大小为 50,交叉概率为 0.9,变异概率为 0.04,进化次数为 500。采样资源处理完所有作业的总时间、资源处理完所有作业的总花费后,更改作业参数,重新创建 50 个作业,并再次测试,共采样 20 次。

根据测试数据,利用式(2)计算出 IAD。对 20 组同一用户参数范围内的作业请求进行测试,得出每个算法在各种用户请求条件下的相应用户综合评价价值,结果如图 3 所示,其中,GA 表示本文算法;cost 和 time 分别表示 GridSim 自带的代价最优算法和时间最优算法。

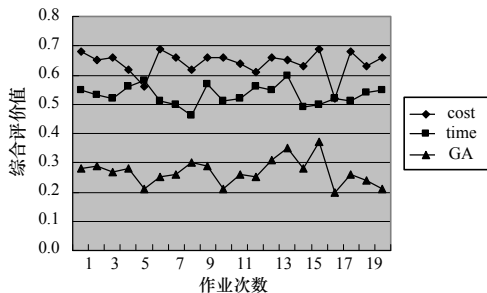


图 3 不同算法的 IAD 值

由图 3 可知,本文算法能比其他 2 种算法得到更好的用户综合评价价值。

测试在本文算法中引入 2 个经济因子后,对作业调度产生的影响。测试时,根据测试环境定义的参数范围创建 50 个作业,采用基于 GA 算法的作业全局调度策略,运行参数设置同上。采样资源处理完所有作业的总时间、资源处理完所有作业的总花费后,对同一组作业设置不同因子并再次测试,共采样 10 组作业。

研究在同一组作业中不同因子情况下,用户完成作业的总时间和总花费,结果如图 4 和图 5 所示,其中,β 是用户作业完成总时间所占的比例因子。由图 5 可知,在相同作业请求下,对于不同因子,当 β 变大、用户作业截止时间要求越高时,作业的执行时间相应变小,且用户代价相应变大。

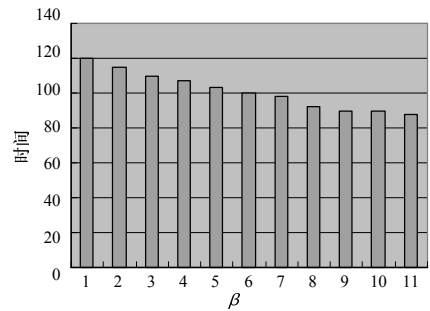


图 4 GA 算法下设置不同因子时的用户作业完成总时间

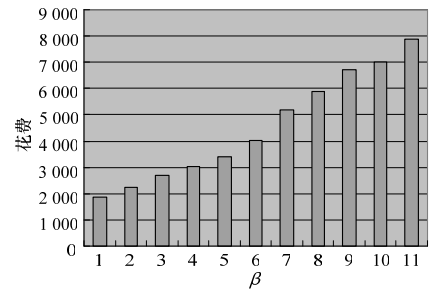


图 5 GA 算法下设置不同因子时的用户总花费

在算法测试过程中发现资源分配不均,部分资源出现空闲。分析后发现,这是由资源的性能差异过大所引起的。通过调整测试环境可以减轻该问题产生的影响。

4 结束语

在本文策略中,用户可以根据不同经济因素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。实验结果证明,基于 GA 的资源调度策略能有效提高用户的总评价价值,从而提高网格系统的利用率和性价比。下一步工作是将此算法移植到真实网格环境中,以验证其性能。

参考文献

- [1] Buyya R, Abramson D, Giddy J. Nimrod/G: An Architecture for a Resource Management and Scheduling System in a Global Computational Grid[C]//Proceedings of the 4th International Conference on High Performance Computing in Asia-Pacific Region. Beijing, China: [s. n.], 2000.
- [2] 穆晓芳,余雪丽,牛瑞萍. 基于拍卖机制的网格在线信誉系统模型[J]. 计算机工程与设计, 2008, 29(4): 979-982.
- [3] 张颖峰,李毓麟. 基于进化算法的网格计算资源管理调度系统[J]. 计算机工程, 2003, 29(15): 110-175.
- [4] 李敏强,寇纪淞,林丹,等. 遗传算法的基本理论与应用[M]. 北京: 科学出版社, 2002.
- [5] Buyya R, Murshed M. GridSim: A Toolkit for the Modeling and Simulation of Distributed Resource Management and Scheduling for Grid Computing[J]. Journal of Concurrency and Computation: Practice and Experience, 2002, 14(13): 1175-1220.

编辑 陈 晖